GLOBAL EDITION

Applied Behavior Analysis

THIRD EDITION

John O. Cooper • Timothy E. Heron • William L. Heward

Applied Behavior Analysis

Third Edition

Global Edition

John O. Cooper

Timothy E. Heron

William L. Heward

All, The Ohio State University

Pearson

Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Dubai • Singapore • Hong Kong
Tokyo • Seoul • Taipei • New Delhi • Cape Town • São Paulo • Mexico City • Madrid • Amsterdam • Munich • Paris • Milan

Director and Publisher: Kevin M. Davis
Portfolio Manager: Drew Bennett
Editor, Global Edition: Punita Kaur Mann
Assistant Project Editor, Global Edition: Aditi Chatterjee
Managing Content Producer: Megan Moffo
Content Producer: Janelle Rogers
Content Producer, Global Edition: Nitin Shankar
Portfolio Management Assistant: Maria Feliberty
Managing Digital Producer: Autumn Benson
Media Producer: Lauren Carlson
Senior Manufacturing Controller, Global Edition: Caterina Pellegrino
Executive Development Editor: Linda Bishop
Executive Field Marketing Manager: Krista Clark
Procurement Specialist: Deidra Headlee
Managing Editor, Media Production, Global Edition: Gargi Banerjee
Manager, Media Production, Global Edition: Vikram Kumar
Cover Design: Lumina Datamatics
Cover Art: Yuliiaart/Shutterstock
Full Service Vendor: SPi Global
Full Service Project Management: Rajakumar Venkatesan, SPi Global

Photo Credits: Jill C. Dardig, p. 4; B. F. Skinner Foundation, p. 25 (left); Julie S. Vargas/B. F. Skinner Foundation, p. 25 (right).

Acknowledgments of third-party content appear on the page within the text, which constitute an extension of this copyright page.

Pearson Education Limited
KAO Two
KAO Park
Harlow
CM17 9SR
United Kingdom

and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsonglobaleditions.com

ISBN 10: 1-292-32463-5
ISBN 13: 978-1-292-32463-0

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

10 9 8 7 6 5 4 3 2 1

Typeset in Times LT Pro by SPi Global
Printed and bound by Vivar in Malaysia

This book is dedicated to Jack Michael, whose extraordinary contributions to behavior analysis will continue to benefit those who teach and practice the science and, most importantly, those whose learning is enhanced by its application.

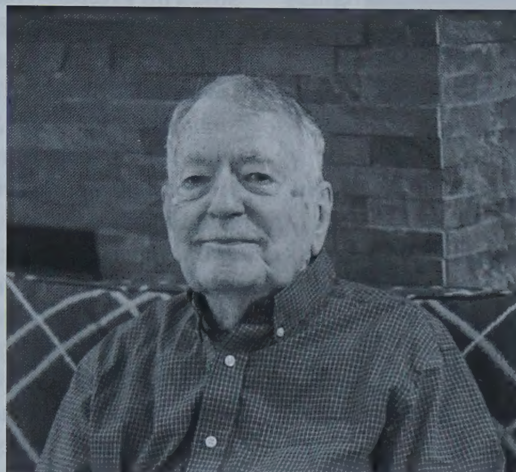


Photo credit: Amber Hutson.

Jack L. Michael

*"One might ask why it is of any value to be able to recognize and correctly name these various effects. I would answer that I have found, for myself at least, that I cannot understand some things unless I can talk about them clearly."**

*From "What Every Student of Behavior Analysis Ought to Learn: A System for Classifying the Multiple Effects of Behavioral Variables," by J. Michael, 1995, *The Behavior Analyst, 18*, p. 284.

ABOUT THE AUTHORS

Photo credit: Jill C. Dardig.

Tim Heron (left), John Cooper (center), and Bill Heward (right)

John Cooper, Tim Heron, and Bill Heward were faculty members at The Ohio State University for a combined 90 years. Together they trained special education classroom teachers and leadership personnel guided by the philosophical, scientific, and technological principles of applied behavior analysis. The Ph.D. program in special education and applied behavior analysis that they and their colleagues developed at OSU was the first doctoral program accredited by the Association for Behavior Analysis International. John, Tim, and Bill each received the Alumni Award for Distinguished Teaching, OSU's highest honor for teaching excellence. They are joint recipients of the Ellen P. Reese Award for Communication of Behavioral Concepts from the Cambridge Center for Behavioral Studies.

John O. Cooper, Ed.D., is Professor Emeritus in the College of Education and Human Ecology at The Ohio State University. His research and teaching interests include precision teaching, inner behavior, fluency building, and verbal behavior. He is a past president of the Standard Celeration Society, past member of the Board of Directors for the Society for the Advancement of Behavior Analysis, and Applied Representative to the Executive Council and as Chairperson, Graduate Program Accreditation Committee of the Association for Behavior Analysis International.

Timothy E. Heron, Ed.D., is Professor Emeritus in the College of Education and Human Ecology at The Ohio State University. Tim's research interests include tutoring systems, inclusion of students with disabilities in general education classrooms, consultation, and self-correction instructional approaches. Tim is co-author of *The Educational Consultant: Helping Professionals, Parents, and Students in Inclusive Classrooms*, Fourth Edition (with Kathleen Harris, 2001). Since 2000, Tim has been an active Federal Aviation Administration, Gold Seal Certified Flight Instructor; has authored the text *Instrument Flying: 10 Indispensable Principles to Know and Remember*; and continues to apply the principles and procedures of applied behavior analysis to the next generation of aviators.

William L. Heward, Ed.D., BCBA-D, is Professor Emeritus in the College of Education and Human Ecology at The Ohio State University. Bill's interests include "low-tech" methods for increasing the effectiveness of group instruction and promoting the generalization and maintenance of newly learned skills. He has authored or co-authored five other books, including *Exceptional Children: An Introduction to Special Education,* Eleventh Edition (with Sheila Alber-Morgan and Moira Konrad, 2017), and *Sign Here: A Contracting Book for Children and Their Parents* (with Jill C. Dardig, 2016). A Fellow and Past President of the Association for Behavior Analysis International, Bill is a recipient of the Fred S. Keller Behavioral Education Award from Division 25 of the American Psychological Association and the Distinguished Psychology Department Alumnus Award from Western Michigan University.

CHAPTER CONTRIBUTORS

Thomas S. Critchfield, Ph.D., is Professor of Psychology at Illinois State University. He earned his doctorate in Behavior Analysis from West Virginia University and completed a post-doctoral fellowship at the Johns Hopkins University School of Medicine. A Fellow and Past President of the Association for Behavior Analysis International and of Division 25 of the American Psychological Association, Tom has served as Associate Editor for the *Journal of the Experimental Analysis of Behavior*, *Perspectives on Behavior Science*, *Mexican Journal of Behavior Analysis*, and *Conch Republic Archives of Psychology*. His interests include derived stimulus relations, punishment and negative reinforcement, effective instruction, and the process of bench-to-bedside scientific translation.

Thomas R. Freeman, M.S., BCBA, is Senior Vice President of ABA Technologies, Inc., where he focuses on the dissemination of behavioral science, in part by helping create curricula and instructional materials for the Florida Institute of Technology ABA Online program. With nearly 40 years of experience in ABA, Tom has held various clinical, supervisory, and administrative positions in Massachusetts and Florida. He has also participated in animal behavior studies with orangutans in Borneo and Hawaiian spinner dolphins, and was Field Director of the University of Hawaii's Humpback Whale Project. Tom is dedicated to applying behavior analysis to mainstream social needs (e.g., general education, environmental issues) and common individual challenges (e.g., anxiety, depression, and grief), and is particularly interested in coordinating ABA and psychiatric services, identifying practices as evidence-based (or not), and studying the evolution of ethics.

Brian A. Iwata, Ph.D., is Distinguished Professor of Psychology and Psychiatry at the University of Florida. He and his students have published over 250 articles and chapters on disorders of learning and behavior and on functional analysis methodology. Brian is a former editor of the *Journal of Applied Behavior Analysis* and past president of the Association for Behavior Analysis International, Division 33 of the American Psychological Association, the Society for the Advancement of Behavior Analysis, the Society for the Experimental Analysis of Behavior, and the Florida Association for Behavior Analysis. He has chaired study sections for both NIH and NIMH and is a fellow of the American Association on Intellectual and Developmental Disabilities, the American Psychological Association, the Association for Behavior Analysis International, and the Association for Psychological Scien ce. In 2015, he received the Gold Medal for Lifetime Achievement in the Application of Psychology from the American Psychological Association.

Linda A. LeBlanc, Ph.D., BCBA-D, Licensed Psychologist, is President of LeBlanc Behavioral Consulting. She received her Ph.D. in 1996 from Louisiana State University. She previously taught at Claremont McKenna College, Western Michigan University, and Auburn University, and was Executive Director of Trumpet Behavioral Health. Her research interests focus on behavioral treatment of autism, technology-based behavioral interventions, behavioral gerontology, supervision and mentorship, and systems development in human services. Dr. LeBlanc has been an associate editor for *Behavior Analysis in Practice, The Analysis of Verbal Behavior,* and the *Journal of Applied Behavior Analysis.* She has been a senior editor for *Education and Treatment of Children* and will serve as the editor of the *Journal of Applied Behavior Analysis* from 2020 to 2022. Linda is the 2016 recipient of the American Psychological Association Nathan H. Azrin Award for Distinguished Contribution in Applied Behavior Analysis, and a fellow of the Association of Behavior Analysis International.

Jose Martinez-Diaz, Ph.D., BCBA-D, is Professor and Director of the School of Behavior Analysis at the Florida Institute of Technology, and CEO of ABA Technologies, Inc., an instructional design and technology company. He earned his doctorate in clinical psychology with an emphasis in behavior analysis and therapy from West Virginia University. Jose's principal areas of interest are practitioner training, professional and ethical issues, instructional design and technology, organizational behavior management, and the conceptual analysis of behavior. A past president of the Florida Association of Behavior Analysis (FABA), Jose has served on the board of directors for the Behavior Analyst Certification Board, the Association of Professional Behavior Analysts (APBA), and the Cambridge Center for Behavioral Studies. Awards recognizing Jose's contributions to effective and ethical practice of behavior analysis include the APBA's Jerry Shook Award and FABA's Charles H. Cox Award for Outstanding Service and Advancement of Behavior Analysis in Florida.

Jack Michael, Ph.D., is Professor Emeritus in the Department of Psychology at Western Michigan University, where he taught for 36 years. His primary scholarly interests are verbal behavior, basic theory regarding motivation, and the technical terminology of behavior analysis. Jack contributed to the founding of the Association for Behavior Analysis International and served as its third president. His publications include the highly acclaimed text *Concepts and Principles of Behavior Analysis* (2004). A Fellow of the Association for Behavior Analysis International and the American Psychological Association, Dr. Michael has received many honors and recognitions, including the Distinguished Service to Behavior Analysis Award from the Association for Behavior Analysis, the 2002 Don Hake Award for Translational Research from Division 25 of the American Psychological Association, the 2012 Victor Laties Lifetime of Service Award from the Society for the Experimental Analysis of Behavior, and Western Michigan University's top two faculty honors: Distinguished Faculty Scholar Award and Distinguished Teaching Award. In 2012, Jack was the first recipient of an award named in his honor by the Verbal Behavior Special Interest Group affiliated with ABAI.

Caio F. Miguel, Ph.D., BCBA-D, is Professor of Psychology at California State University, Sacramento. His research interests span basic, applied, and conceptual issues in the study of motivation, verbal behavior, covert mediation, and derived stimulus relations. Caio has served as editor of *The Analysis of Verbal Behavior* and associate editor for the *Journal of Applied Behavior Analysis*. His publications have appeared in English, Portuguese, and Spanish language journals, and he has given hundreds of professional presentations throughout North America, South America, and Europe. Caio received the 2013–2014 Award for Outstanding Scholarly Work from the College of Social Sciences and Interdisciplinary Studies at CSU Sacramento, and the 2014 Outstanding Mentor Award by the Student Committee of the Association for Behavior Analysis International.

Nancy A. Neef, Ph.D., is Professor Emeritus in the College of Education and Human Ecology at The Ohio State University. She has served as editor of the *Journal of Applied Behavior Analysis,* as president of the Society for the Experimental Analysis of Behavior, and on the Executive Council and as chair of the publication board for the Association of Behavior Analysis International. Nancy has published more than 60 articles and chapters in the areas of developmental disabilities, research methodology, and instructional technology. Much of her research has focused on extensions and applications of basic research in the assessment and treatment of attention-deficit hyperactivity disorder. Nancy was the recipient of the first Distinguished Alumnus Achievement Award in Psychology from Western Michigan University and the 2006 Award for Outstanding Research in Applied Behavior Analysis from Division 25 of the American Psychological Association.

Stephanie M. Peterson, Ph.D., BCBA-D, is Professor and Chair of the Department of Psychology at Western Michigan University. Her primary research interests are choice and concurrent schedules of reinforcement in the treatment of severe problem behavior and in the functional analysis of problem behavior. Stephanie also has interests in applications of behavior analysis to educational interventions and teacher training. She has served on the editorial boards of the *Journal of Applied Behavior Analysis* and *The Behavior Analyst* and currently serves as a consulting senior editor for *Education and Treatment of Children*. She is a former member of the Board of Directors for the Behavior Analyst Certification Board.

Carol Pilgrim, Ph.D., is Professor of Psychology at the University of North Carolina, Wilmington. Her primary research interests are in the analysis, application, and conceptual treatment of relational stimulus control, particularly stimulus equivalence. Carol is a former editor of *The Behavior Analyst* and associate editor of the *Journal of the Experimental Analysis of Behavior* and *The Behavior Analyst.* She has served as President of the Association for Behavior Analysis International (ABAI), the Society for the Advancement of Behavior Analysis, Division 25 of the American Psychological Association (APA), and the Southeastern Association for Behavior Analysis. Carol is a fellow of ABAI and Division 25 of APA, and she has been honored with the North Carolina Board of Governors Teaching Excellence Award (2003), the UNCW Faculty Scholarship Award (2000) and Graduate Mentor Award (2008), and the ABAI Student Committee Outstanding Mentor Award (2006) and Distinguished Service to Behavior Analysis Award (2017).

Ruth Anne Rehfeldt, Ph.D., BCBA-D, is Professor of Behavior Analysis and Therapy at Southern Illinois University. She completed her Ph.D. at the University of Nevada. Dr. Rehfeldt has published over 100 articles and book chapters in behavior analysis, primarily on basic and applied investigations of verbal behavior and derived relational responding, relational frame theory, and acceptance and commitment therapy. Ruth Anne served as the editor and business manager for *The Psychological Record* for 12 years. She is or has been an editorial board member for a number of behavior analytic journals, including *Journal of Applied Behavior Analysis, Journal of the Experimental Analysis of Behavior,* and *The Analysis of Verbal Behavior*. She has held a number of leadership positions within the Association for Behavior Analysis International. Dr. Rehfeldt has won a number of teaching and research awards during her tenure at Southern Illinois University.

Richard G. Smith, Ph.D., BCBA-D, LBA-TX, is Associate Professor in the Department of Behavior Analysis at the University of North Texas, where he served as Department Chair for 13 years. Rick received his master's and doctoral degrees at the University of Florida. His primary research interest is in the assessment and treatment of behavior disorders in persons with developmental disabilities, with specific areas of focus in motivational variables and advances in functional analysis procedures. A former associate editor for the *Journal of Applied Behavior Analysis,* Rick's work has been recognized with the American Psychological Association's Division 25 B. F. Skinner Award for Innovative and Important Research by a New Researcher (1997), the Texas Chapter of the American Association on Intellectual and Developmental Disabilities' Research Award (2000), the Texas Association for Behavior Analysis' Career Contributions to Behavior Analysis in Texas Award (2014), and the University of North Texas President's Special Faculty Recognition Award (2017).

Mark L. Sundberg, Ph.D., BCBA-D, is a Licensed Psychologist in private practice. He specializes in language research and the development of language assessment and intervention programs for children and adults with language delays. Mark is founder and past editor of the journal *The Analysis of Verbal Behavior,* a past president of The Northern California Association for Behavior Analysis, and a past chair of the Publication Board of the Association for Behavior Analysis International, and has served on the Board of Directors of the B. F. Skinner Foundation. Mark is the author of *The Verbal Behavior Milestones Assessment and Placement Program (VB-MAPP),* and co-author (with James W. Partington) of the books *Teaching Language to Children with Autism or Other Developmental Disabilities* and the original *Assessment of Basic Language and Learning Skills: The ABLLS.* Mark has received several awards, including the 2001 Distinguished Psychology Department Alumnus Award from Western Michigan University and the 2013 Jack Michael Outstanding Contributions in Verbal Behavior Award from ABAI's Verbal Behavior Special Interest Group.

PREFACE

As it was 17 years ago, when we began writing the previous edition of this book, our overriding goal for the third edition was to produce an accurate, comprehensive, and contemporary description of applied behavior analysis. The result is a text that demands concentrated and serious study.

Despite its size, scope, and in-depth treatment of concepts, principles, procedures, and issues, *Applied Behavior Analysis, Third Edition*, should be viewed as an introductory text for two reasons. First, the reader need not possess any specialized prerequisite knowledge to understand the content. Second, attaining a full understanding of applied behavior analysis requires considerable study and guided experience beyond this text. There is no topic presented within these pages that has not been treated in greater depth elsewhere. Students of applied behavior analysis should build upon what they learn from this book by reading other sources. How much reading is needed to fully grasp and appreciate applied behavior analysis? Don Baer (2005), one of the co-founders of applied behavior analysis, estimated that

> [T]he cost of knowing well the basic principles and paradigms of the theoretical and experimental aspects of behavior analysis would require about 2,000 pages and some laboratory experience. ABA shares the same basic principles with the theoretical and experimental branches of behavior analysis and adds to them an even larger number of secondary principles, strategies, and tactics for making those basic principles work in the real world as they do in the laboratory. ABA also adds a set of principles about ethical and humane practice, prominent among which is the need to be certain, through constant and extensive measurement and experimentation, that the particular case in hand is going well and will continue to go well—because it will change as it progresses. The cost of knowing all that is, I estimate, about 3,000 pages of reading and several years of supervised practical experience. (pp. 27–28)

The field has grown markedly since its formal inception in 1968, so much so that Baer's 3000-page reading assignment might now exceed 4000 pages, or more. We trust that this book's 800-plus pages will contribute to Baer's reading assignment for many future behavior analysts. Specific suggestions for additional readings in applied behavior analysis and in the conceptual and basic research branches of behavior analysis are cited throughout this text.

Again, while our objective is to provide a complete description of the principles and procedures for changing and analyzing socially important behavior, mastery of this book's content represents the beginning, not the end, of one's study of applied behavior analysis. If our efforts as textbook writers and chapter contributors, combined with those of instructors who assign this book, are successful, the dedicated student will come away with a sound repertoire of fundamental knowledge about applied behavior analysis. In turn, this knowledge will serve as the foundation for more advanced study and supervised practica that ultimately will lead to independent efforts to change and understand behavior that are scientifically sound, socially significant, and ethically appropriate.

TERMINOLOGY

Meaningful description of any scientific activity necessitates a standard set of technical terms. Effectively communicating the design, implementation, outcomes, and/or theoretical bases of an applied behavior analysis requires the accurate use of the discipline's terminology. Throughout this text, we have made every effort to define and use behavior analytic terminology in a conceptually systematic and consistent manner. Mastering the specialized vocabulary of applied behavior analysis is an important initial step in embracing the science and participating effectively as a researcher, practitioner, or consumer. We encourage students to study the field's terminology with diligence. Toward that end, the third edition includes a glossary of more than 500 technical terms and concepts.

REFERENCES, EXTRACTS, NOTES, AND GRAPHS

An important function of any introductory text to a scientific discipline is to expose students to the empirical and conceptual literature of that field. This edition contains more than 2700 citations to primary-source publications, including historically important experiments (e.g., the first graph presented by B. F. Skinner in his 1938 book *The Behavior of Organisms*), and classic and contemporary examples of applied behavior analysis research—most of which were published in the field's flagship journal, the *Journal of Applied Behavior Analysis.* We also make extensive use of quotations and extracts from key publications representing the conceptual literature. We have done this not only for the historical and/or technical authority these authors provide, but also because their inclusion increases students' exposure to and appreciation for the field's rich primary-source literature.

The third edition includes more than 150 graphs of data from peer-reviewed research, many of which are accompanied by detailed descriptions of the study's methodology. We have a fourfold purpose for providing many procedures, graphs, and references. First, we want to illustrate behavior analysis principles and procedures with actual applications and real data, not hypothetical examples. Second, reading the procedural descriptions will help students appreciate the high degree of technical precision and control of complex environments that researchers and practitioners must achieve to solve problems and show functional relations between variables. Third, the references provide students whose interests are piqued by the descriptions or graphs with directions to the original studies for more in-depth study. Finally, the graphs provide multiple opportunities for students to develop and refine—through practice and discussion with their instructors, mentors, and fellow students—higher-level visual analysis skills.

THIRD EDITION CONTENT ENHANCEMENTS AND FEATURES

Applied behavior analysis has become more mature and sophisticated since the second edition was published. Although the basic principles of behavior remain unchanged, advances in all three interrelated domains of the science of behavior—theoretical, basic research, and applied research—have improved our understanding of those principles and led to increased effectiveness in developing and applying effective and humane behavior change interventions. These developments are reflected in the more than 1,000 new references to the conceptual, basic, and applied literatures of behavior analysis added to this edition.

Chapters by Outstanding Authors

The third edition includes seven chapters authored by prominent scholars in applied behavior analysis. This group of contributors includes the current and two former editors of the *Journal of Applied Behavior Analysis,* two previous editors of *The Analysis of Verbal Behavior,* and associate editors of the *Journal of the Experimental Analysis of Behavior*. Members of this well-known and prolific group of behavior analysts first reported some of the most significant advances in behavior analysis in publications.

Negative Reinforcement

In Chapter 12, Negative Reinforcement, Rick Smith and Brian Iwata present an authoritative account of this commonly misunderstood and misapplied form of reinforcement. In addition to precisely defining this principle, dispelling misconceptions about it, and illustrating applications across a broad spectrum of cases, Smith and Iwata provide specific guidelines for incorporating negative reinforcement into behavior change interventions.

Motivation

Until recently, motivation, a major topic in psychological theories and everyday explanations of behavior, has been an assumed, but inadequately understood, topic in behavior analysis. Due largely to the work of Jack Michael, behavior analysts now have a much better understanding of motivation and its role within applied behavior analysis. In Chapter 16, Motivating Operations, Jack Michael and Caio Miguel explain how certain antecedent events have dual motivating effects: a behavior-altering effect, which makes certain behaviors more (or less) likely; and a value-altering effect, which makes certain events more (or less) effective as reinforcement.

Verbal Behavior

In Chapter 18, Verbal Behavior, Mark Sundberg contrasts B. F. Skinner's functional analysis of verbal behavior with traditional approaches to language, defines and gives examples of basic types of elementary verbal operants (e.g., mands, tacts, intraverbals), and describes implications and applications for these concepts in designing and implementing language intervention programs.

Equivalence-based Instruction

In Chapter 19, Equivalence-based Instruction, Carol Pilgrim builds upon Sidman's groundbreaking research on stimulus equivalence to explain the conditions in which learners acquire new skills and verbal relations without direct instruction on those skills. Pilgrim defines equivalence-based instruction, describes its key outcomes—class formation, delayed emergence, class expansion and class merger, transfer of function, and contextual control—and shows how lessons can be designed to promote those outcomes.

Nonequivalence Relations

In Chapter 20, Engineering Emergent Learning with Nonequivalence Relations, Tom Critchfield and Ruth Anne Rehfeldt explain how people make sense of and function effectively in a world of arbitrary relations in which stimuli "go together," not because they share physical properties, but rather because social-verbal reinforcement contingencies teach people to relate them in a certain way. Relational frame theory (RFT) and acceptance and commitment therapy (ACT), a therapeutic approach grounded in RFT, are described.

Functional Behavior Assessment

In Chapter 27, Functional Behavior Assessment, Stephanie Peterson and Nancy Neef describe one of the more significant developments in applied behavior analysis. Functional behavior assessment has become a well-established method for discovering the function that a problem behavior serves for a person (e.g., to obtain social attention, to avoid an assigned task, to provide sensory stimulation), information that enables practitioners to design interventions that teach adaptive replacement behaviors that serve the same function.

Ethics

In Chapter 31, Ethical and Professional Responsibilities of Applied Behavior Analysts, Tom Freeman, Linda LeBlanc, and Jose Martinez-Diaz clarify what ethical behavior is, explain why ethical behavior is a necessary part of the applied behavior analyst's repertoire, review ethical codes of conduct for behavior analysts, and describe specific procedures for ensuring and assessing ethical practice. New content regarding client services (e.g., informed consent, conflict of interest) and, importantly, ethical implications of new technologies, social media, and professional networking to support ethical behavior is presented.

TEXT ORGANIZATION AND STRUCTURE

The book's 31 chapters are organized into 13 parts. The two chapters in Part 1 describe some tenets that are fundamental to all scientific endeavors, outline a history of behavior analysis as a natural science approach to understanding behavior, define applied behavior analysis, and describe principles and concepts of that science. Parts 2 and 3 examine the elements necessary for an applied behavior analysis. Part 2 presents considerations, criteria, and procedures for selecting, defining, and measuring *applied behavior.* The five chapters in Part 3 examine the logic

and operation of specific tactics for the experimental *analysis* of behavior–environment relations, and some issues in planning, replicating, and evaluating analyses of behavior.

The seven chapters in Parts 4 through 6 explore the two most significant principles of behavior—reinforcement and punishment; how antecedent events alter one's motivation to respond; and how behavior comes under the discriminative control of environmental conditions. Part 7 is a detailed examination of B. F. Skinner's analysis of verbal behavior and its implications and applications for language development. The five chapters in Part 8 describe how applied behavior analysts use equivalence-based instruction, nonequivalence relations, imitation, modeling, observational learning, shaping, and chaining to develop new skills and patterns of behavior from simple to complex.

Part 9 details how problem behaviors can be decreased with nonpunishment interventions: extinction, differential reinforcement, and antecedent interventions. Part 10 describes functional behavioral assessment, sophisticated methods for determining the purpose that problem behavior serves for a person, and important information that leads to the design of treatments that replace the problem behavior with adaptive alternative behavior serving the same function.

Part 11 describes four special applications of behavior change technology: token economy, contingency contracting, group contingencies, and self-management. Part 12 outlines strategies and tactics for increasing the likelihood that efforts to change behavior yield generalized outcomes: behavior changes that maintain across time, occur in appropriate settings and situations beyond the training setting, and spread to other useful behaviors. The book's final part describes ethical and professional responsibilities of behavior analysts, ethical implications of new technologies, social media, and professional networking.

SUPPLEMENTS AND RESOURCES FOR STUDENTS AND INSTRUCTORS

Instructor's Resource Manual and Test Bank (ISBN 1-292-32464-3)

An Instructor's Resource Manual includes suggestions for learning activities, additional Experiencing Firsthand exercises, supplementary lectures, case study analyses, discussion topics, group activities, additional media resources, and answers to all multiple-choice questions and essay-type questions. The *Test Bank* that accompanies this text contains more multiple-choice and essay-type questions. Some items (lower-level questions) simply ask students to identify or explain concepts and principles they have learned. But many others (higher-level questions) ask students to apply those same concepts and principles to specific classroom situations—that is, to actual student behaviors and teaching strategies.

Powerpoint® Slides (ISBN 1-292-32466-X)

The PowerPoint slides include key concept summarizations, diagrams, and other graphic aids to enhance learning. They are designed to help students understand, organize, and remember core concepts and theories.

Companion Website

Text Content Related to the Behavior Analyst Certification Board® BCBA® & BCABA® Behavior Analyst Task List© Fifth Edition

The Behavior Analyst Certification Board® (BACB®) is a nonprofit corporation established in 1998 to meet professional credentialing needs identified by behavior analysts, governments, and consumers of behavior analysis services. To be certified as a Board Certified Behavior Analyst® (BCBA®) or a Board Certified Assistant Behavior Analyst® (BCaBA®), a person must meet academic-degree, educational, and practical-experience eligibility requirements and then pass a psychometrically sound examination. The BCBA and BCaBA examinations are based on the *BCBA/BCaBA Task List* (5th ed.; BACB, 2017a), which was developed by 16 subject matter experts and subsequently content validated through a survey of more than 6000 BACB certificants (BACB, 2017b). The complete *BCBA/BCaBA Task List* (5th ed.) is available on the companion website (www.pearsonglobaleditions.com).

We have connected the content of this text to the tasks that the BACB determined are necessary to function as an entry-level behavior analyst. A chart identifying which Task List items are covered in each chapter is also provided on the companion website. Due to the complex nature of applied behavior analysis, in which the concepts and principles and their application are interrelated and not easily or effectively presented in a linear fashion, some Task List items are covered in more than one chapter. Students studying for BCBA and BCaBA examinations can look up key words from Task List items in the Subject Index to identify the page numbers where relevant information about each item in the *BCBA/BCaBA Task List* (5th ed.) can be found.

This text presents the basic knowledge that a qualified behavior analyst must possess. Although mastering this content will help you obtain a passing score on the BCBA or BCaBA examinations, two important qualifiers must be recognized. First, the BCBA and BCaBA examinations require knowledge beyond that included in this, or any, single textbook. Therefore, to further prepare for the examinations we encourage students to study original sources, engage in supervised practica, and discuss areas of personal interest with trusted and competent mentors. Second, no matter how accurate, extensive, and current this textbook may be, and no matter how thoroughly a student masters its content, he or she will not be fully qualified to function as a behavior analyst. Successful completion of the required coursework in behavior analysis is but one step in the preparation to become a BCBA or a BCaBA. For the most recent information on the BACB requirements, visit the Behavior Analyst Certification Board's website at www.BACB.com.

Behavior Analyst Certification Board. (2017a). *BCBA/BCaBA task list (5th ed.).* Littleton, CO: Author.

Behavior Analyst Certification Board. (2017b, January). *BACB newsletter.* https://www.bacb.com/wp-content/uploads/170113-newsletter.pdf

Acknowledgments

The third edition of *Applied Behavior Analysis* is the product of the collective and cumulative efforts of many people. Although space limitations prevent us from thanking everyone by name, we would like to acknowledge those who made substantial contributions to the book's content or production during the 3 years we spent revising the text. First and foremost, we are deeply grateful to the authors of the seven contributed chapters: Rick Smith and Brian Iwata (Negative Reinforcement); Jack Michael and Caio Miguel (Motivating Operations); Mark Sundberg (Verbal Behavior); Carol Pilgrim (Equivalence-based Instruction); Tom Critchfield and Ruth Anne Rehfeldt (Engineering Emergent Learning with Nonequivalence Relations); Stephanie Peterson and Nancy Neef (Functional Behavior Assessment); and Tom Freeman, Linda LeBlanc, and Jose Martinez-Diaz (Ethical and Professional Responsibilities of Applied Behavior Analysts). Because of their efforts, third edition readers will be introduced to important topics in applied behavior analysis by scholars whose research has helped to define and develop those areas.

We thank the Behavior Analyst Certification Board (BACB) for allowing us to integrate the Behavior Analyst Certification Board® BCBA/BCaBA Task List (5th ed.) throughout the revised edition of our text. We are especially grateful to Jim Carr, Chief Executive Officer of the Behavior Analyst Certification Board. Jim graciously continued the arrangement with the BACB we first developed with Jerry Shook for the second edition.

In-depth reviews of the second edition by Jan Montgomery (Florida Institute of Technology), Bethany Raiff (Rowan University), Corey Robertson (Florida Institute of Technology), Tina Sidener (Caldwell University), and Carlos A. Zuluaga (Florida Institute of Technology) offered helpful recommendations for improvement that we incorporated into this edition. The third edition is also enhanced by input and assistance from the following colleagues: Bill Ahearn (New England Center for Children), Sheila Alber-Morgan (The Ohio State University), Nikolay Alipov (Russian National Investigation Medical University), Fernando Armendariz (FABA, Inc.), Judah Axe (Simmons College), Vince Carbone (The Carbone Clinic), Sharon Chien (SEEK), Darlene Crone-Todd (Salem University), Alyce Dickinson (Western Michigan University), Anca Dumitrescu (ATCA, Romania), Ronnie Dietrich (Wing Institute), Tim Hackenberg (Reed College), Louis Hagopian (Kennedy Krieger Institute), Rob Holdsambeck (Cambridge Center for Behavioral Studies), Mickey Keenan (Ulster University, Northern Ireland), Tracy Kettering (Bancroft), Jonathan Kimball (private practice, Maine), Moira Konrad (The Ohio State University), Douglas Kostewicz (University of Pittsburgh), Rick Kubina (Penn State University), Ed Morris (University of Kansas), Dave Palmer (Smith College), Robert Ross (Beacon ABA Services), Hank Schlinger (California State University, Los Angeles), Majda Seuss (Association for Behavior Analysis International), and Janet Twyman (Center for Innovations in Learning).

Turning more than 2500 pages of manuscript into the book you are holding required the support and contributions of a talented team of publishing professionals at Pearson. Copy editor Joanne "Bonnie" Boehme's talent in turning rough prose into clear and understandable text is evident on every page. The bumpy spots that remain are our responsibility. Production Editor Clara Bartunek and Content Producer Janelle Rogers worked closely and patiently with us throughout. Ann Castel Davis, our former editor, was instrumental in convincing us to tackle this edition. We are grateful to Ann and to Director and Publisher Kevin Davis, for their long-standing commitment not only to this title but also to publishing books on behavior analysis and evidence-based practices in special education.

Last, but certainly not least, among those we wish to acknowledge for important and continuing contributions to the "White Book" is Keith "Dutch" Van Norman. Dutch designed the creative and captivating covers of the second and third editions. Thanks, Dutch!

Throughout our careers, each of us has benefited immensely from mentors, colleagues, and students who provided us with the instruction, models, and inspiration needed to attempt to write a book like *Applied Behavior Analysis*. To the extent that our goals in writing this book are achieved, each of them will have played an important role. To the professors who first taught us about applied behavior analysis—Saul Axelrod, Vance Cotter, Todd Eachus, Dick Malott, Jack Michael, Joe Spradlin, and Don Whaley—we will be indebted always. Many faculty colleagues at Ohio State helped us create and sustain an academic environment where work to advance, teach, and apply behavior analysis for the betterment of learners with disabilities is valued. Their individual and collective support is appreciated. We'd be remiss not to single out our longtime department chair, Thomas M. Stephens, for his sustained support and leadership. Although not formerly trained as a behavior analyst, Tom was deeply committed to evidence-based instruction and direct and frequent measurement of student performance. Tom provided much-needed administrative backing, sometimes in the form of covering fire that allowed us to build an applied behavior analysis program in a college dominated by nonbehavioral perspectives. We also wish to thank the many students whom we have been lucky enough to teach and learn with over the years. Your enthusiasm and dedication inspired, encouraged, and motivated us.

Finally, we want to acknowledge the support of our families, past and present: Bunny, Chris, Sharon, Greg, Brian, and Carroll and Vera Cooper; Marge Heron, Kathy, Patrick, and Leia Rogers, Christine, Matt, Audrey and Brady Harsh, and Ray and Bernice Heron; and Jill Dardig, Lee Heward, Lynn Heward and Marcus Bonde, and Joe and Helen Heward. Without their enduring love and support, we would never have attempted the first edition of this book, let alone a second and a third!

BRIEF CONTENTS

CONTENTS

GLOBAL EDITION ACKNOWLEDGMENTS

Pearson would like to thank Ruchika Kanwal for her work on this Global Edition.

PART 1

Introduction and Basic Concepts

We believe that prior to learning specific principles and procedures for analyzing and changing behavior, the student of applied behavior analysis should be introduced to the historical and conceptual foundations of the science. Basic knowledge and appreciation of the scientific and philosophical underpinnings of behavior analysis are requisites to a thorough understanding of the discipline's nature, scope, and potential. We also believe a preliminary overview of basic concepts, principles, and terminology makes the in-depth study of behavior analysis to follow more effective. The two chapters in Part 1 support these two beliefs. Chapter 1 describes the scientific, conceptual, and philosophical roots of applied behavior analysis and identifies the discipline's defining dimensions, characteristics, and overall goals. Chapter 2 defines the field's fundamental elements—behavior and the environmental antecedent and consequential events that influence it—and introduces key terms and principles that describe relationships among these elements.

CHAPTER 1

Definition and Characteristics of Applied Behavior Analysis

LEARNING OBJECTIVES

- Describe the basic characteristics and goals of science.
- Explain behavior in accordance with the philosophical assumptions of behavior analysis.
- Explain determinism as it relates to behavior analysis.
- State distinguishing features of mentalistic and environmental explanations of behavior.
- Describe and explain behavior in behavior analytic terms.
- State and describe each of the dimensions of applied behavior analysis.

[S]ince I was a child I always found my biggest reinforcer was something called understanding. I liked to know how things worked. And of all of the things in the world there are to understand, it became clear to me that the most fascinating was what people do. I started with the usual physical science stuff, and it was intriguing to me to understand how radios work, and how electricity works, and how clocks work, etcetera. But when it became clear to me that we could also learn how people work—not just biologically, but behaviorally—I thought that's the best of all. Surely, everyone must agree that that's the most fascinating subject matter. That there could be a science of behavior, of what we do, of who we are? How could you resist that?

—Donald M. Baer in Heward & Wood, (2003, p. 302)

Applied behavior analysis is a science devoted to understanding and improving human behavior. Other disciplines have similar intents. What sets applied behavior analysis apart? The answer lies in its focus, goals, and methods. Applied behavior analysts focus on behaviors of social importance, they intervene with research-based strategies and tactics to improve the targeted behaviors, and they use scientific methods—objective description, measurement, and experimentation—to demonstrate reliable relations between their interventions and the behavioral improvements. In short, applied behavior analysis, or ABA, is a scientific approach for discovering environmental variables that reliably influence socially significant behavior and for developing a technology of behavior change that takes practical advantage of those discoveries.

This chapter briefly outlines the history and development of behavior analysis, discusses the philosophy that underlies the science, and identifies defining dimensions and characteristics of applied behavior analysis. Because applied behavior analysis is first and foremost a science, we begin with an overview of precepts shared by scientists in all disciplines.

SCIENCE: BASIC CHARACTERISTICS AND A DEFINITION

Science is a systematic approach for seeking and organizing knowledge about the natural world. Before offering a definition of science, we discuss the purpose of science and the basic assumptions and attitudes that guide the work of all scientists, irrespective of their fields of study.

Purpose of Science

The overall goal of science is to achieve a thorough understanding of the phenomena under study—socially important behavior change, in the case of applied behavior analysis. Science differs from other sources of knowledge or ways we obtain knowledge about the world around us (e.g., contemplation, common sense, logic, authority figures, religious or spiritual beliefs, political campaigns, advertisements, testimonials). Science seeks to discover nature's truths: facts and universal laws that exist and operate independent of the opinions and beliefs of any person or group, including the scientist. Therefore, scientific knowledge must be separated from any personal, political, economic, or other reasons for which it was sought. Although it is frequently misused, science is not a tool for validating the cherished or preferred versions of "the truth" held by any group, corporation, government, or institution.

Different types of scientific investigations yield knowledge enabling one or more of three levels of understanding: description, prediction, and control. Each level of understanding contributes to the scientific knowledge base of a given field of inquiry.

Description

Systematic observation enhances the understanding of a given phenomenon by enabling scientists to describe it accurately. Descriptive knowledge consists of a collection of facts about the observed events that can be quantified, classified, and examined for possible relations with other known facts—a necessary and important activity for any scientific discipline. The knowledge obtained from descriptive studies often suggests possible hypotheses or questions for additional research.

The work of John James Audubon, a naturalist and painter in the early 19th century, provides a classic example of descriptive science. While observing birds in their natural habitat, Audubon documented their habits with extensive field notes and made detailed drawings. He identified 25 new species of birds. His major work, *The Birds of America* (Audubon, 1827–1838), contains 435 hand-colored life-sized prints of birds in their natural habitat and is considered one of the finest ornithological works ever completed.

White's (1975) study of classroom teachers' "natural rates" of approval (verbal praise or encouragement) and disapproval (criticisms, reproach) is an example of descriptive research in applied behavior analysis. Observations of 104 classroom teachers in grades 1 to 12 yielded two major findings: (a) Rates of teacher praise dropped with each grade level, and (b) in every grade after second, teachers delivered statements of disapproval to students at rates exceeding their rates of praise. The results of this descriptive study led to dozens of subsequent studies aimed at discovering factors responsible for the disappointing findings, analyzing the effects of disproportionate rates of disapproval and praise on student behavior, and increasing teachers' effective use of praise (e.g., Alber, Heward, & Hippler, 1999; Duchaine, Jolivette, & Fredrick, 2011; Fullerton, Conroy, & Correa, 2009; Mrachko, Kostewicz, & Martin, 2017; Niwayama & Tanaka-Matsumi, 2016; Sutherland, Wehby, & Yoder, 2002).

Prediction

A second level of scientific understanding occurs when repeated observations reveal that two events consistently covary with each other. That is, in the presence of one event (e.g., approaching winter) another event occurs (or fails to occur) with some specified probability (e.g., certain birds fly south). When systematic covariation between two events is found, this relationship—termed a *correlation*—can be used to predict the relative probability that one event will occur, based on the presence of the other event. "We obviously cannot intervene or manipulate the movement of the stars or planets, but by studying their movements we can gauge the seasons and when we can plant crops to produce a bountiful harvest" (Moore, 2010, p. 48).

Because no variables are manipulated or controlled by the researcher, a correlational study cannot demonstrate whether one of the observed variables is responsible for the changes in the other variable, and no such relations should be inferred. A strong correlation exists between hot weather and an increased incidence of drowning deaths, but we should not assume that a hot and humid day causes anyone to drown. Hot weather also correlates with other factors, such as an increased number of people (both swimmers and nonswimmers) seeking relief in the water, and many instances of drowning have been found to be a function of factors such as the use of alcohol or drugs, the relative swimming skills of the victims, strong riptides, and the absence of supervision by lifeguards.[1]

In addition to their usefulness in aiding prediction, the findings of correlational studies can suggest the possibility of causal relations, which can then be explored with experimental studies. The most common type of correlational study reported in the applied behavior analysis literature compares the relative rates or conditional probabilities of two or more observed (but not manipulated) variables (e.g., Atwater & Morris, 1988; Symons, Hoch, Dahl, & McComas, 2003; Thompson & Iwata, 2001). For example, McKerchar and Thompson (2004) found correlations between problem behavior exhibited by 14 preschool children and the following consequent events: teacher attention (100% of the children), presentation of some material or item to the child (79% of the children), and escape from instructional tasks (33% of the children). The results of this study not only provide empirical validation for the social consequences typically used in clinical settings to analyze the variables maintaining children's problem behavior, but also increase confidence in the prediction that interventions based on the findings from such assessments will be relevant to the conditions that occur naturally in preschool classrooms (see Chapter 27). In addition, by revealing the high probabilities with which teachers responded to problem behavior in ways that are likely to maintain and strengthen it, McKerchar and Thompson's findings also point to the need to train teachers in more effective ways to respond to problem behavior.

Control

The ability to predict with a certain degree of confidence is a valuable and useful result of science; prediction enables preparation. However, the greatest potential benefits from science are derived from the third, and highest, level of scientific understanding—control. Evidence of the kinds of control that can be derived from scientific findings in the physical and biological sciences surrounds us in the everyday technologies we take for granted: pasteurized milk and the refrigerators we store it in; flu shots and the automobiles we drive to go get them; pain relievers and the televisions that bombard us with advertisements and news stories about the drugs.

> The scientific "system," like the law, is designed to enable us to handle a subject matter more efficiently . . . When we have discovered the laws which govern a part of the world about us, we are then ready to deal effectively with that part of the world. By predicting the occurrence of an event we are able to prepare for it. By arranging conditions in ways specified by the laws of a system, we not only predict, we control: we "cause" an event to occur or to assume certain characteristics. (Skinner, 1953, pp. 13–14)

Functional relations, the primary products of basic and applied research in behavior analysis, provide the kind

of scientific understanding that is most valuable and useful to the development of a technology for changing behavior. A **functional relation** exists when a well-controlled experiment demonstrates that a specific change in one event (the *dependent variable*) is reliably produced by specific manipulations of another event (the *independent variable*), and that the change in the dependent variable was unlikely to be the result of other extraneous factors (*confounding variables*).

Johnston and Pennypacker (1980) described functional relations as "the ultimate product of a natural scientific investigation of the relation between behavior and its determining variables" (p. 16).

> Such a "co-relation" is expressed as $y = f(x)$, where x is the independent variable or argument of the function, and y is the dependent variable. In order to determine if an observed relation is truly functional, it is necessary to demonstrate the operation of the values of x in isolation and show that they are sufficient for the production of y. . . . [H]owever, a more powerful relation exists if necessity can be shown (that y occurs *only if* x occurs). The most complete and elegant form of empirical inquiry involves applying the experimental method to identifying functional relations. (Johnston & Pennypacker, 1993a, p. 239)

The understanding gained by the scientific discovery of functional relations is the basis of applied technologies in all fields.

Assumptions and Attitudes of Science

> *Science is first of all a set of attitudes.*
>
> —B. F. Skinner, (1953, p. 12)

The definition of science lies not in test tubes, spectrometers, or electron accelerators, but in the behavior of scientists. To begin to understand any science, we need to look past the apparatus and instrumentation that are most readily apparent and examine what scientists do.[2] The pursuit of knowledge is properly called *science* when it is carried out according to general methodological precepts and expectations that define science. All scientists share a fundamental assumption about the nature of events that are amenable to investigation by science, general notions about basic strategy, and perspectives on how to view their findings. These attitudes of science—determinism, empiricism, experimentation, replication, parsimony, and philosophic doubt—constitute a set of overriding assumptions and values that guide the work of all scientists (Whaley & Surratt, 1968).

Determinism

Science is predicated on the assumption of **determinism**. All scientists presume that the universe is a lawful and orderly place in which all phenomena occur as the result of other events. In other words, events do not just happen willy-nilly; they are related in systematic ways to other factors, which are themselves physical phenomena amenable to scientific investigation.

If the universe were governed by *accidentalism,* a philosophical position antithetical to determinism that holds that events occur by accident or without cause, or by *fatalism,* the belief that events are predetermined, the scientific discovery of functional relations and use of those discoveries to improve things would be impossible.

> If we are to use the methods of science in the field of human affairs, we must assume behavior is lawful and determined. We must expect to discover what a man does is the result of specifiable conditions and that once these conditions have been discovered, we can anticipate and to some extent determine his actions. (Skinner, 1953, p. 6)

Determinism plays a pivotal dual role in the conduct of scientific practice: It is at once a philosophical stance that does not lend itself to proof and the confirmation that is sought by each experiment. In other words, the scientist first assumes lawfulness and then proceeds to look for lawful relations (Delprato & Midgley, 1992).

Empiricism

> *When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind*
>
> — Lord Kelvin, (1824–1907)

Scientific knowledge is built on, above all, **empiricism**—the practice of objective observation and measurement of the phenomena of interest. Objectivity in this sense means "independent of the individual prejudices, tastes, and private opinions of the scientist. Results of empirical methods are objective in that they are open to anyone's observation and do not depend on the subjective belief of the individual scientist" (Zuriff, 1985, p. 9).

In the prescientific era (and in nonscientific and pseudoscientific activities today) (Nichols, 2017), knowledge was (and is) the product of contemplation, speculation, personal opinion, authority, and the "obvious" logic of common sense. The scientist's empirical attitude, however, demands objective observation based on thorough description, systematic and repeated measurement, and precise quantification of the phenomena of interest.

As it is in every scientific field, empiricism is the foremost rule in behavior analysis. Every effort to understand, predict, and improve behavior hinges on the behavior analyst's ability to completely define, systematically observe, and accurately and reliably measure occurrences and nonoccurrences of the behavior of interest.

Experimentation

Experimentation is the basic strategy of most sciences. Whaley and Surratt (1968) used the following anecdote to introduce the need for experimentation.

> A man who lived in a suburban dwelling area was surprised one evening to see his neighbor bow to the four winds, chant a strange melody, and dance around his front lawn beating a small drum. After witnessing the same ritual for over a month, the man became overwhelmed with curiosity and decided to look into the matter.

> "Why do you go through this same ritual each evening?" the man asked his neighbor.
>
> "It keeps my house safe from tigers," the neighbor replied.
>
> "Good grief!" the man said. "Don't you know there isn't a tiger within a thousand miles of here?"
>
> "Yeah," the neighbor smiled. "Sure works, doesn't it!" (pp. 23–2 to 23–3)

When events are observed to covary or occur in close temporal sequence, a functional relation may exist, but other factors may be responsible for the observed values of the dependent variable. To investigate the possible existence of a functional relation, an experiment (or better, a series of experiments) must be performed in which the factor(s) suspected of having causal status are systematically controlled and manipulated while the effects on the event under study are carefully observed.

Reliably predicting and controlling any phenomena, including the presence of tigers in one's backyard, requires identifying and manipulating the factors that influence those phenomena. One way that the individual described previously could use the experimental method to evaluate the effectiveness of his ritual would be to first move to a neighborhood in which tigers are regularly observed and then systematically manipulate the use of his anti-tiger ritual (e.g., 1 week off, 1 week on, 1 week off, 1 week on) while observing and recording the presence of tigers under the no-ritual and ritual conditions.

> The experimental method is a method for isolating the relevant variables within a pattern of events. . . . [W]hen the experimental method is employed, it is possible to change one factor at a time (independent variable) while leaving all other aspects of the situation the same, and then to observe what effect this change has on the target behavior (dependent variable). Ideally, a functional relation may be obtained. Formal techniques of experimental control are designed to make sure that the conditions being compared are otherwise the same. Use of the experimental method serves as a necessary condition (sine qua non) to distinguish the experimental analysis of behavior from other methods of investigation. (Dinsmoor, 2003, p. 152)

Thus, an **experiment** is a controlled comparison of some measure of the phenomenon of interest (the dependent variable) under two or more different conditions in which only one factor at a time (the independent variable) differs from one condition to another. Strategies and tactics for conducting experiments in applied behavior analysis are described in Chapters 7 through 10.

Most of the studies cited in this text are experiments that have demonstrated or discovered a functional relation between a target behavior and one or more environmental variables. Such studies are said to have achieved a functional analysis. The term **functional analysis** has two meanings in contemporary behavior analysis literature. In its original and most fundamental usage, *functional analysis* denotes demonstrations of functional relations between environmental variables and behavior. Schlinger and Normand (2013) reported that Skinner used the term 36 times in *Science and Human Behavior* and cited this example:

> The external variables of which behavior is a function provide for what may be called a causal or *functional analysis*. We undertake to predict and control the behavior of the individual organism. This is our "dependent variable"—the effect for which we are to find the cause. Our "independent variables"—the causes of behavior—are the external conditions of which behavior is a function. Relations between the two—the "cause-and-effect relationships" in behavior—are the laws of a science. (Skinner, 1953, p. 35, italics added)

Iwata, Dorsey, Slifer, Bauman, and Richman (1982) introduced the second and today most widely recognized usage of *functional analysis* in their groundbreaking article describing an experimental methodology for determining environmental variables and contingencies maintaining problem behavior (see Chapter 27). In its original meaning, functional analysis provides the very foundation for an experimental science of behavior; as a method for assessing the controlling variables for problem behavior, functional analysis informs the design of effective treatments.

Replication

The results of a single experiment—no matter how well it was designed and conducted, no matter how clear and impressive the findings—are never sufficient to earn an accepted place among the scientific knowledge base of any field. Although the data from a single experiment have value in their own right and cannot be discounted, only after an experiment has been replicated a number of times with the same basic pattern of results are scientists convinced of the findings.

Replication—repeating of experiments (as well as repeating independent variable conditions within experiments)—"pervades every nook and cranny of the experimental method" (Johnston & Pennypacker, 1993a, p. 244). Replication is the primary method with which scientists determine the reliability and usefulness of their findings and discover their mistakes (Johnston & Pennypacker, 1980; 1993a; Sidman, 1960). Replication—not the infallibility or inherent honesty of scientists—is the primary reason science is a self-correcting enterprise that ultimately gets it right (Skinner, 1953).

How many times must an experiment be repeated with the same results before the scientific community accepts the findings? There is no required number of replications, but the greater the importance of the findings to theory or practice, the greater the number of replications to be conducted. Chapters 7 through 10 explain the role of replication in behavioral research and describe replication strategies used by applied behavior analysts.

Parsimony

One dictionary definition of *parsimony* is great frugality, and in a special way this connotation accurately describes the behavior of scientists. As an attitude of science, **parsimony** requires

that all simple, logical explanations for the phenomenon under investigation be ruled out, experimentally or conceptually, before more complex or abstract explanations are considered. Parsimonious interpretations help scientists assess and fit new findings within the field's existing knowledge base. A fully parsimonious interpretation consists only of those elements that are necessary and sufficient to explain the phenomenon at hand. The attitude of parsimony is so critical to scientific explanations that it is sometimes referred to as the Law of Parsimony (Whaley & Surratt, 1968), a "law" derived from *Occam's Razor,* credited to William of Occam (c. 1285–1349), who stated: "One should not increase, beyond what is necessary, the number of entities required to explain anything." In other words, given a choice between two competing and compelling explanations for the same phenomenon, one should shave off extraneous variables and choose the simplest explanation, the one that requires the fewest assumptions.

Philosophic Doubt

The attitude of **philosophic doubt** requires the scientist to continually question the truthfulness of what is regarded as fact. Scientific knowledge must always be viewed as tentative. Scientists must be willing to set aside their most cherished beliefs and findings and replace them with the knowledge derived from new discoveries.

Good scientists maintain a healthy level of skepticism. Although being skeptical of others' research may be easy, a more difficult but critical characteristic of scientists is that they remain open to the possibility—as well as look for evidence—that their own findings or interpretations are wrong. "Science is a willingness to accept facts even when they are opposed to wishes" (Skinner, 1953, p. 12). As Oliver Cromwell (1650) stated in another context: "I beseech you . . . think it possible you may be mistaken." For the true scientist, "new findings are not problems; they are opportunities for further investigation and expanded understanding" (Todd & Morris, 1993, p. 1159).

Practitioners should be as skeptical as researchers. The skeptical practitioner not only requires scientific evidence before implementing a new practice, but also evaluates continually its effectiveness once the practice has been implemented. Practitioners must be particularly skeptical of extraordinary claims made for the effectiveness of new theories, therapies, or treatments (Foxx & Mulick, 2016; Maurice, 2017).

> Claims that sound too good to be true usually are. Extraordinary claims require extraordinary evidence (Sagan, 1996; Shermer, 2002). What constitutes extraordinary evidence? In the strictest sense, and the sense that should be employed when evaluating claims of educational effectiveness, evidence is the outcome of the application of the scientific method to test the effectiveness of a claim, a theory, or a practice. The more rigorously the test is conducted, the more often the test is replicated, the more extensively the test is corroborated, the more extraordinary the evidence. Evidence becomes extraordinary when it is extraordinarily well tested. (Silvestri & Heward, 2016, p. 149)

We end our discussion of philosophic doubt with two pieces of advice, one from Carl Sagan and one from B. F. Skinner: "The question is not whether we *like* the conclusion that emerges out of a train of reasoning, but whether the conclusion *follows* from the premise or starting point and whether that premise is true" (Sagan, 1996, p. 210). "Regard no practice as immutable. Change and be ready to change again. Accept no eternal verity. Experiment" (Skinner, 1979, p. 346).

Other Important Attitudes and Values

The six attitudes of science that we have examined are necessary features of science and provide an important context for understanding applied behavior analysis.

However, the behavior of most productive and successful scientists is also characterized by qualities such as thoroughness, curiosity, perseverance, diligence, ethics, and honesty. Scientists acquire these traits because behaving in such ways has proven beneficial to the progress of science.

A Definition of Science

Science has no universally accepted, standard definition. We offer the following definition as one that encompasses the previously discussed purposes and attitudes of science, irrespective of the subject matter. **Science** is a systematic approach to understanding natural phenomena—as evidenced by description, prediction, and control—that relies on determinism as its fundamental assumption, empiricism as its prime directive, experimentation as its basic strategy, replication as its necessary requirement for believability, parsimony as its conservative value, and philosophic doubt as its guiding conscience.

A BRIEF HISTORY OF BEHAVIOR ANALYSIS

The science of behavior analysis entails three interrelated domains: philosophy, basic research, and applied research. **Behaviorism** is the philosophy of the science of behavior, basic research is the province of the experimental analysis of behavior (EAB), and developing a technology for improving behavior is the concern of applied behavior analysis (ABA). To be fully understood, applied behavior analysis must be considered in the context of the philosophy and basic research traditions and findings from which it evolved and remains connected today. This section provides an elementary description of the basic tenets of behaviorism and outlines some of the major events that have marked the development of behavior analysis.[3] Table 1.1 lists major books, journals, and professional organizations that have contributed to the advancement of behavior analysis since the 1930s.

Watson's Stimulus–Response Behaviorism

Psychology in the early 1900s was dominated by the study of states of consciousness, images, and other mental processes. Introspection, the act of carefully observing one's own conscious thoughts and feelings, was a primary method of investigation.

TABLE 1.1 Books, Journals, and Organizations That Have Played a Major Role in the Development and Dissemination of Behavior Analysis

Decade	Books	Journals	Organizations
1930s	*The Behavior of Organisms*—Skinner (1938)	*The Psychological Record* (1937)	
1940s	*Walden Two*—Skinner (1948)		
1950s	*Principles of Psychology*—Keller and Schoenfeld (1950)	*Journal of the Experimental Analysis of Behavior* (1958)	Society for the Experimental Analysis of Behavior (SEAB) (1957)
	Science and Human Behavior—Skinner (1953)		
	Schedules of Reinforcement—Ferster and Skinner (1957)		
	Verbal Behavior—Skinner (1957)		
1960s	*Tactics of Scientific Research*—Sidman (1960)	*Journal of Applied Behavior Analysis (1968)*	American Psychological Association's Division 25 Experimental Analysis of Behavior (1964)
	Child Development, Vols. I & II—Bijou and Baer (1961, 1965)		Experimental Analysis of Behaviour Group (UK) (1965)
	The Analysis of Behavior—Holland and Skinner (1961)		
	Research in Behavior Modification—Krasner and Ullmann (1965)		
	Operant Behavior: Areas of Research and Application—Honig (1966)		
	The Analysis of Human Operant Behavior—Reese (1966)		
	Principles of Behavioral Analysis—Millenson (1967)		
	Behavior Principles—Ferster and Perrott (1968)		
	Contingencies of Reinforcement: A Theoretical Analysis—Skinner (1969)		
1970s	*Beyond Freedom and Dignity*—Skinner (1971)	*Behaviorism* (1972) (became *Behavior and Philosophy* in 1990)	Norwegian Association for Behavior Analysis (1973)
	Elementary Principles of Behavior—Whaley and Malott (1971)	*Revista Mexicana de Analisis de la Conducta* (1975)	Midwestern Association for Behavior Analysis (MABA) (1974)
	About Behaviorism—Skinner (1974)	*Behavioural Processes* (1976)	Mexican Society of Behavior Analysis (1975)
	Single Case Experimental Designs—Hersen and Barlow (1976)	*Behavior Modification* (1977)	Association for Behavior Analysis (formerly, MABA) (1978)
	Applying Behavior-Analysis Procedures with Children and Youth—Sulzer-Azaroff and Mayer (1977)	*Journal of Organizational Behavior Management* (1977)	
	Learning—Catania (1979)	*Education & Treatment of Children* (1977)	
		The Behavior Analyst (1978)	
1980s	*Strategies and Tactics of Human Behavioral Research*—Johnston and Pennypacker (1980)	*Journal of Precision Teaching and Celeration* (formerly, *Journal of Precision Teaching*) (1980)	Society for the Advancement of Behavior Analysis (1980)
	Behaviorism: A Conceptual Reconstruction—Zuriff (1985)	*Analysis of Verbal Behavior* (1982)	
	Recent Issues in the Analysis of Behavior—Skinner (1989)	*Behavioral Interventions* (1986)	Cambridge Center for Behavioral Studies (1981)

(continued)

TABLE 1.1 (continued)

Decade	Books	Journals	Organizations
		Japanese Journal of Behavior Analysis (1986)	Japanese Association for Behavior Analysis (1983)
		Behavior Analysis Digest (1989)	
		Behavioural Pharmacology (1989)	
1990s	*Concepts and Principles of Behavior Analysis*—Michael (1993)	*Behavior and Social Issues* (1991)	Accreditation of Training Programs in Behavior Analysis (Association for Behavior Analysis) (1993)
	Understanding Behaviorism: Science, Behavior, and Culture—Baum (1994)	*Journal of Behavioral Education* (1991)	Behavior Analyst Certification Board (BACB) (1998)
	Radical Behaviorism: The Philosophy and the Science—Chiesa (1994)	*Journal of Positive Behavior Interventions* (1999)	Council of Directors of Graduate Programs in Behavior Analysis (Association for Behavior Analysis) (1999)
	Equivalence Relations and Behavior—Sidman (1994)	*The Behavior Analyst Today* (1999)	First Board Certified Behavior Analysts (BCBA) credentialed by the BACB (1999)
	Behavior Analysis and Learning—Pierce and Epling (1995)		
	Functional Analysis of Problem Behavior—Repp and Horner (1999)		
2000s	*Relational Frame Theory: A Post-Skinnerian Account of Human Language and Cognition*—Hayes, Barnes-Holmes, and Roche (2001)	*European Journal of Behavior Analysis* (2000)	
	Conceptual Foundations of Radical Behaviorism—Moore (2008)	*Behavioral Development Bulletin* (2002)	
		Journal of Early and Intensive Behavior Intervention (2004)	European Association for Behaviour Analysis (2002)
		Brazilian Journal of Behavior Analysis (2005)	Association for Professional Behavior Analysts (APBA) (2007)
		International Journal of Behavioral Consultation and Therapy (2005)	Association for Behavior Analysis International (ABAI) (formerly, ABA) (2008)
2010s	*Handbook of Applied Behavior Analysis*—Fisher, Piazza, and Roane (2011)	*Behavior Analysis in Practice* (2011)	First Registered Behavior Technician (RBT) credentialed by the BACB (2014)
	The Science of Consequences—Schneider (2012)	*Journal of Contextual Behavioral Science* (2012)	BACB credentials the 30,000th behavior analyst (2018)
	APA Handbook of Behavior Analysis—Madden (2013)	*Operants* (2014)	Membership in ABAI and affiliate chapters surpasses 26,000 in 63 countries (2018)
	Radical Behaviorism for ABA Practitioners—Johnston (2013)	*Behavior Analysis: Research and Practice* (formerly, *The Behavior Analyst Today* (2015)	
	The Wiley-Blackwell Handbook of Operant and Classical Conditioning—McSweeney and Murphy (2014)	*Perspectives on Behavior Science* (formerly, *The Behavior Analyst*) (2018)	
	The Nurture Effect: How the Science of Human Behavior Can Improve Our Lives & Our World—Biglan (2015)		

Note: Books are listed by initial year of publication. Some titles are available in more recent editions


B. F. Skinner *(left)* in his Indiana University lab circa 1945 and *(right)* circa 1967.

Although the authors of several texts in the first decade of the 20th century defined psychology as the science of behavior (see Kazdin, 1978), John B. Watson is widely recognized as the spokesman for a new direction in the field of psychology. In his influential article "Psychology as the Behaviorist Views It," Watson (1913) wrote:

> Psychology as the behaviorist views it is a purely objective experimental branch of natural science. Its theoretical goal is the prediction and control of behavior. Introspection forms no essential part of its methods, nor is the scientific value of its data dependent upon the readiness with which they lend themselves to interpretation in terms of consciousness. (p. 158)

Watson argued that the proper subject matter for psychology was not states of mind or mental processes but observable behavior. Further, the objective study of behavior as a natural science should consist of direct observation of the relationships between environmental stimuli (S) and the responses (R) they evoke. Watsonian behaviorism became known as stimulus–response (S–R) psychology. Although scientific evidence was insufficient to support S–R psychology as a workable explanation for most behavior, Watson was confident that his new behaviorism would lead to the prediction and control of human behavior and that it would allow practitioners to improve performance in areas such as education, business, and law. Watson (1924) made bold claims concerning human behavior, as illustrated in this famous quotation:

> Give me a dozen healthy infants, well-formed, and my own specified world to bring them up in and I'll guarantee to take any one at random and train him to become any type of specialist I might select—doctor, lawyer, artist, merchant-chief and, yes, even beggar-man and thief, regardless of his talents, penchants, tendencies, abilities, vocations, and race of his ancestors. I am going beyond my facts and I admit it, but so have the advocates of the contrary and they have been doing it for many thousands of years. (p. 104)

It is unfortunate that such extraordinary claims were made, exaggerating the ability to predict and control human behavior beyond the scientific knowledge available. The quotation just cited has been used to discredit Watson and continues to be used to discredit behaviorism in general, even though the behaviorism that underlies contemporary behavior analysis is fundamentally different from the S–R paradigm. Nevertheless, Watson's contributions were of great significance: He made a strong case for the study of behavior as a natural science on a par with the physical and biological sciences.[4]

Experimental Analysis of Behavior

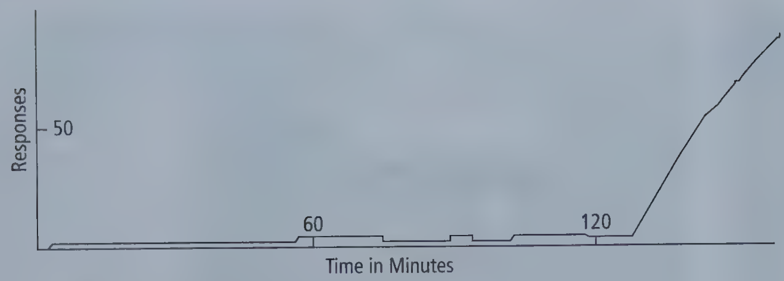
> *[Science] is a search for order. It begins, as we all begin, by observing single episodes, but it quickly passes on to the general rule, to scientific law.*
>
> —B. F. Skinner, (1953, pp. 13–14)

> *I had the clue from Pavlov: Control your conditions and you will see order.*
>
> —B. F. Skinner, (1956, p. 223)

The experimental branch of behavior analysis formally began with the publication of B. F. Skinner's *the Behavior of Organisms* (1938). The book summarized Skinner's laboratory research conducted from 1930 to 1937 and brought into perspective two kinds of behavior: respondent and operant.

Respondent behavior is reflexive behavior as in the tradition of Ivan Pavlov (1927). Respondents are elicited, or "brought out," by stimuli that immediately precede them. The antecedent stimulus (e.g., bright light) and the response it elicits (e.g., pupil constriction) form a functional unit called a *reflex.* Respondent behaviors are essentially involuntary and occur whenever the eliciting stimulus is presented.

Skinner was "interested in giving a scientific account of all behavior, including that which Descartes had set aside as 'willed' and outside the reach of science" (Glenn, Ellis, & Greenspoon, 1992, p. 1330). But, like other psychologists of the time, Skinner found that the S–R paradigm could not explain a great deal of behavior, particularly behaviors that had no



Original Conditioning
All responses to the lever were reinforced. The first three reinforcements were apparently ineffective. The fourth is followed by a rapid increase in rate.

Figure 1.1 The first data set in B. F. Skinner's *The Behavior of Organisms: An Experimental Analysis* (1938).

Based on *the Behavior of Organisms: An Experimental Analysis* by B. F. Skinner, p. 67. Original copyright 1938 by Appleton-Century. Copyright 1991 by B. F. Skinner Foundation, Cambridge, MA. Used by permission.

apparent antecedent causes in the environment. Compared to reflexive behavior with its clear eliciting events, much of the behavior of organisms appeared spontaneous or "voluntary." In an attempt to explain the mechanisms responsible for "voluntary" behavior, other psychologists postulated mediating variables inside the organism in the form of hypothetical constructs such as cognitive processes, drives, and free will. Skinner took a different tack. Instead of creating **hypothetical constructs**, presumed but unobserved entities that could not be manipulated in an experiment, Skinner continued to look in the environment for the determinants of behavior that did not have apparent antecedent causes.

> He did not deny that physiological variables played a role in determining behavior. He merely felt that this was the domain of other disciplines, and for his part, remained committed to assessing the causal role of the environment. This decision meant looking elsewhere in time. Through painstaking research, Skinner accumulated significant, if counterintuitive, evidence that behavior is changed less by the stimuli that precede it (though context is important) and more by the consequences that immediately follow it (i.e., consequences that are contingent upon it). The essential formulation for this notion is S–R–S, otherwise known as the three–term contingency. It did not replace the S–R model—we still salivate, for instance, if we smell food cooking when we are hungry. It did, however, account for how the environment "selects" the great part of learned behavior.
>
> With the three-term contingency Skinner gave us a new paradigm. He achieved something no less profound for the study of behavior and learning than Bohr's model of the atom or Mendel's model of the gene. (Kimball, 2002, p. 71)

Skinner called the second type of behavior *operant behavior.*[5] Operant behaviors are not elicited by preceding stimuli but instead are influenced by stimulus changes that have followed the behavior in the past. Skinner's most powerful and fundamental contribution to our understanding of behavior was his discovery and experimental analyses of the effects of consequences on behavior. The operant three-term contingency as the primary unit of analysis was a revolutionary conceptual breakthrough.

Skinner (1938) argued that the analysis of operant behavior "with its unique relation to the environment presents a separate important field of investigation" (p. 438). He named this new science the **experimental analysis of behavior** and outlined the methodology for its practice. Simply put, Skinner recorded the rate at which a single subject (he initially used rats and later, pigeons) emitted a given behavior in a controlled and standardized experimental chamber.

The first set of data Skinner presented in *The Behavior of Organisms* was a graph that "gives a record of the resulting change in behavior" (p. 67) when a food pellet was delivered immediately after a rat pressed a lever (see Figure 1.1). Skinner noted that the first three times that food followed a response "had no observable effect" but that "the fourth response was followed by an appreciable increase in rate showing a swift acceleration to a maximum" (pp. 67–68).

Skinner's investigative procedures evolved into an elegant experimental approach that enabled clear and powerful demonstrations of orderly and reliable functional relations between behavior and various types of environmental events.[6] By systematically manipulating the arrangement and scheduling of stimuli that preceded and followed behavior in literally thousands of laboratory experiments from the 1930s through the 1950s, Skinner and his colleagues and students discovered and verified the basic principles of operant behavior that continue to provide the empirical foundation for behavior analysis today. Description of these principles of behavior—general statements of functional relations between behavior and environmental events—and tactics for changing behavior derived from those principles constitute a major portion of this text.

Skinner's Radical Behaviorism

> *Behavior analysts dispense with the myth of the inner person as creator of behavior. Both philosophically and empirically to the behavior analyst, we are what we do, and when and where we do it.*
>
> —Murray Sidman, (2013, p. xvi)

In addition to being the founder of the experimental analysis of behavior, B. F. Skinner wrote extensively on the philosophy of that science. Without question, Skinner's writings have been the most influential both in guiding the practice of the science of behavior and in proposing the application of the principles of behavior to new areas.[7] In 1948 Skinner published *Walden Two,* a fictional account of how the philosophy and principles of behavior might be used in a utopian community (see Altus & Morris, 2009). This was

followed by his classic text, *Science and Human Behavior* (1953), in which he speculated on how the principles of behavior might be applied to complex human behavior in areas such as education, religion, government, law, and psychotherapy.

Much of Skinner's writing was devoted to the development and explanation of his philosophy of behaviorism. Skinner began his book *About Behaviorism* (1974) with these words:

> Behaviorism is not the science of human behavior; it is the philosophy of that science. Some of the questions it asks are these: Is such a science really possible? Can it account for every aspect of human behavior? What methods can it use? Are its laws as valid as those of physics and biology? Will it lead to a technology, and if so, what role will it play in human affairs? (p. 1)

The behaviorism Skinner pioneered differed significantly (indeed, radically) from other psychological theories, including other forms of behaviorism. Although there were, and remain today, many psychological models and approaches to the study of behavior, **mentalism** is the common denominator among most.

> In general terms, *mentalism* may be defined as an approach to the study of behavior which assumes that a mental or "inner" dimension exists that differs from a behavioral dimension. This dimension is ordinarily referred to in terms of its neural, psychic, spiritual, subjective, conceptual, or hypothetical properties. Mentalism further assumes that phenomena in this dimension either directly cause or at least mediate some forms of behavior, if not all. These phenomena are typically designated as some sort of act, state, mechanism, process, or entity that is causal in the sense of initiating or originating. Mentalism regards concerns about the origin of these phenomena as incidental at best. Finally, mentalism holds that an adequate causal explanation of behavior must appeal directly to the efficacy of these mental phenomena. (Moore, 2003, pp. 181–182)

Hypothetical constructs and explanatory fictions are the stock and trade of mentalism, which has dominated Western intellectual thought and most psychological theories (Descartes, Freud, Piaget), and it continues to do so into the 21st century. Freud, for example, created a complex mental world of hypothetical constructs—the id, ego, and superego—that he contended were key to understanding a person's actions.

Hypothetical constructs—"theoretical terms that refer to a possibly existing, but at the moment unobserved process or entity" (Moore, 1995, p. 36)—can be neither observed nor experimentally manipulated (MacCorquodale & Meehl, 1948; Zuriff, 1985). Free will, readiness, innate releasers, language acquisition devices, storage and retrieval mechanisms for memory, and information processing are all examples of hypothetical constructs that are inferred from behavior. Although Skinner (1953, 1974) clearly indicated that it is a mistake to rule out events that influence our behavior because they are not accessible to others, he believed that using presumed but unobserved mentalistic fictions (i.e., hypothetical constructs) to explain the causes of behavior contributed nothing to a functional account.

Consider a typical laboratory situation. A food-deprived rat pushes a lever each time a light comes on and receives food, but the rat seldom pushes the lever when the light is off (and if it does, no food is delivered). When asked to explain why the rat pushes the lever only when the light is on, most will say that the rat has "made the association" between the light being on and food being delivered when the lever is pressed. As a result of making that association, the animal now "knows" to press the lever only when the light is on. Attributing the rat's behavior to a hypothetical cognitive process such as associating or to something called "knowledge" adds nothing to a functional account of the situation. First, the environment (in this case, the experimenter) paired the light and food availability for lever presses, not the rat. Second, the knowledge or other cognitive process that is said to explain the observed behavior is itself unexplained, which begs for still more conjecture.

The "knowledge" that is said to account for the rat's performance is an example of an **explanatory fiction**, a fictitious variable that often is simply another name for the observed behavior that contributes nothing to an understanding of the variables responsible for developing or maintaining the behavior. Explanatory fictions are the key ingredient in "a circular way of viewing the cause and effect of a situation" (Heron, Tincani, Peterson, & Miller, 2005, p. 274) that lead to a false sense of understanding.

> Turning from observed behavior to a fanciful inner world continues unabated. Sometimes it is little more than a linguistic practice. We tend to make nouns of adjectives and verbs and must then find a place for the things the nouns are said to represent. We say that a rope is strong and before long we are speaking of its strength. We call a particular kind of strength tensile, and then explain that the rope is strong *because* it possesses tensile strength. The mistake is less obvious but more troublesome when matters are more complex.
>
> Consider now a behavioral parallel. When a person has been subject to mildly punishing consequences in walking on a slippery surface, he may walk in a manner we describe as cautious. It is then easy to say that he walks with caution or that he shows caution. There is no harm in this until we begin to say that he walks carefully *because* of his caution. (Skinner, 1974, pp. 165–166, emphasis added)

It is widely believed that Skinner rejected all events that could not be independently verified by agreement among observers. However, Skinner was explicit early on that he valued effective action over agreement among observers.

> The ultimate criterion for the goodness of a concept is not whether two people are brought into agreement but whether the scientist who uses the concept can operate successfully upon his material—all by himself if need be. What matters to Robinson Crusoe is not whether he is agreeing with himself but whether he is getting anywhere with his control over nature. (Skinner, 1945, p. 293).

Pragmatism, the philosophical position that "the truth value of a statement is a function of how well the statement promotes effective action" (Moore, 2008, p. 400), continues to be a primary criterion by which behavior analysts judge the value of their findings (Leigland, 2010; Moxley, 2004).[8]

In reality, there are many kinds of behaviorism—structuralism, methodological behaviorism, and forms of behaviorism that use cognitions as causal factors (e.g., cognitive behavior modification and social learning theory)—in addition to the radical behaviorism of Skinner. Structuralism and methodological behaviorism do reject all events that are not operationally defined by objective assessment. Structuralists avoid mentalism by restricting their activities to descriptions of behavior. They make no scientific manipulations; accordingly, they do not address questions of causal factors. Methodological behaviorists differ from the structuralists by using scientific manipulations to search for functional relations between events. Uncomfortable with basing their science on unobservable phenomena, some early behaviorists either denied the existence of "inner variables" or considered them outside the realm of a scientific account. Such an orientation is often referred to as **methodological behaviorism**.

Methodological behaviorists acknowledge the existence of mental events but do not consider them in the analysis of behavior (Skinner, 1974). Methodological behaviorists' reliance on public events, excluding private events, restricts the knowledge base of human behavior and discourages innovation in the science of behavior. Methodological behaviorism is restrictive because it ignores areas of major importance for an understanding of behavior.

Contrary to another common misconception, Skinner did not object to cognitive psychology's concern with private events (i.e., events taking place "inside the skin") (Moore, 2000). Skinner was the first behaviorist to view thoughts and feelings (he called them "private events") as behavior to be analyzed with the same conceptual and experimental tools used to analyze publicly observable behavior, not as phenomena or variables that exist within and operate according to principles of a separate mental world. "I contend that my toothache is just as physical as my typewriter" (Skinner, 1945, p. 294).

Essentially, Skinner's behaviorism makes three major assumptions regarding the nature of private events: (a) Private events such as thoughts and feelings are behavior; (b) behavior that takes place within the skin is distinguished from other ("public") behavior only by its inaccessibility; and (c) private behavior is influenced by (i.e., is a function of) the same kinds of variables as publicly accessible behavior.

> We need not suppose that events which take place within an organism's skin have special properties for that reason. A private event may be distinguished by its limited accessibility but not, so far as we know, by any special structure of nature. (Skinner, 1953, p. 257)

By incorporating private events into an overall conceptual system of behavior, Skinner created a **radical behaviorism** that includes and seeks to understand all human behavior. "What is inside the skin, and how do we know about it? The answer is, I believe, the heart of radical behaviorism" (Skinner, 1974, p. 218). The proper connotations of the word *radical* in radical behaviorism are *far-reaching* and *thoroughgoing,* connoting the philosophy's inclusion of all behavior, public and private. *Radical* is also an appropriate modifier for Skinner's form of behaviorism because it represents a dramatic departure from other conceptual systems in calling for

> probably the most drastic change ever proposed in our way of thinking about man. It is almost literally a matter of turning the explanation of behavior inside out. (Skinner, 1974, p. 256)

Skinner and the philosophy of radical behaviorism acknowledge the events on which fictions such as cognitive processes are based. Radical behaviorism does not restrict the science of behavior to phenomena that can be detected by more than one person. In the context of radical behaviorism, the term *observe* implies "coming into contact with" (Moore, 1984). Radical behaviorists consider private events such as thinking or sensing the stimuli produced by a damaged tooth to be no different from public events such as oral reading or sensing the sounds produced by a musical instrument. According to Skinner (1974), "What is felt or introspectively observed is not some nonphysical world of consciousness, mind, or mental life but the observer's own body" (pp. 18–19).

The analysis of private events is a major aspect of radical behaviorism and indispensable for a comprehensive science of behavior (Palmer, 2011). Moore (1980, 2015) stated it concisely:

> For radical behaviorism, private events are those events wherein individuals respond with respect to certain stimuli accessible to themselves alone. . . . The responses that are made to those stimuli may themselves be public, i.e., observable by others, or they may be private, i.e., accessible only to the individual involved. Nonetheless, to paraphrase Skinner (1953), it need not be supposed that events taking place within the skin have any special properties for that reason alone. . . . For radical behaviorism, then, one's responses with respect to private stimuli are equally lawful and alike in kind to one's responses with respect to public stimuli. (1980, p. 460)

> [T]hese events are critical to understanding behavior in all its complexity. Just as importantly, they need not be formulated in different terms and with different concepts that are publicly observable behavior events. (2015, p. 18)

Scientists and practitioners are affected by their own social context, and institutions and schools are dominated by mentalism (Heward & Cooper, 1992; Kimball, 2002). A firm grasp of the philosophy of radical behaviorism, in addition to knowledge of principles of behavior, can help the scientist and practitioner resist the mentalistic approach of dropping the search for controlling variables in the environment and drifting toward explanatory fictions in the effort to understand behavior. The principles of behavior and the procedures presented in this text apply equally to public and private events. Radical

behaviorism is the philosophical position underlying the content presented in this text.

As Friman (2017) noted, Skinner's behaviorism viewed behavior as a natural science.

> By taking this stand, he was promoting a larger idea, specifically that behavior was solely a physical phenomenon brought about, maintained, strengthened, or weakened solely by physical (environmental) events. In other words, he was promoting the idea that behavior is a function of environmental circumstances and their context. This is the most powerful idea ever invented by mankind for understanding, knowing, and approaching human behavior especially when it is a problem. (p. 176)

A thorough discussion of radical behaviorism is far beyond the scope of this text. The serious student of applied behavior analysis will devote considerable study to Skinner's original writings and to other authors who have critiqued, analyzed, and extended the philosophical foundations of the science of behavior.[9] (See Box 1.1 for Don Baer's perspectives on the meaning and importance of radical behaviorism.)

Applied Behavior Analysis

The first study to report the human application of principles of operant behavior was conducted by Fuller (1949). The subject was an 18-year-old boy with profound developmental disabilities who was described in the language of the time as a "vegetative idiot." He lay on his back, unable to roll over. Fuller filled a syringe with a warm sugar-milk solution and injected a small amount of the fluid into the young man's mouth every time he moved his right arm (that arm was chosen because he moved it infrequently). Within four sessions the boy was moving his arm to a vertical position at a rate of three times per minute.[10]

> The attending physicians . . . thought it was impossible for him to learn anything—according to them, he had not learned anything in the 18 years of his life—yet in four experimental sessions, by using the operant conditioning technique, an addition was made to his behavior which, at this level, could be termed appreciable. Those who participated in or observed the experiment are of the opinion that if time permitted, other responses could be conditioned and discriminations learned. (Fuller, 1949, p. 590)

During the 1950s and into the early 1960s researchers used the methods of the experimental analysis of behavior to determine whether the principles of behavior demonstrated in the laboratory with nonhuman subjects could be replicated with humans. According to Thompson and Hackenberg (2009), "the field of applied analysis emerged from the experimental analysis of behavior, like Adam's rib" (p. 271).

Much of the early research with human subjects was conducted in clinic or laboratory settings. Although the participants typically benefited from these studies by learning new behaviors, the researchers' major purpose was to determine whether the basic principles of behavior discovered in the laboratory operated with humans. For example, Sidney Bijou (1955, 1957, 1958)[11] researched several principles of behavior with typically developing subjects and people with intellectual disabilities; Don Baer (1960, 1961, 1962) examined the effects of punishment, escape, and avoidance contingencies on preschool children; and Ogden Lindsley (1956; Lindsley & Skinner, 1954) assessed the effects of operant conditioning on the behavior of adults with schizophrenia. These early researchers clearly established that the principles of behavior are applicable to human behavior, and they set the stage for the later development of applied behavior analysis.

The branch of behavior analysis that would later be called applied behavior analysis (ABA) can be traced to the 1959 publication of Ayllon and Michael's paper titled "The Psychiatric Nurse as a Behavioral Engineer." The authors described how direct care personnel in a state hospital used a variety of techniques based on the principles of behavior to improve the functioning of residents with psychotic disorders or intellectual disabilities. During the 1960s many researchers began to apply principles of behavior in an effort to improve socially important behavior, but these early pioneers faced many problems. Laboratory techniques for measuring behavior and for controlling and manipulating variables were sometimes unavailable, or their use was inappropriate in applied settings. As a result, the early practitioners of applied behavior analysis had to develop new experimental procedures as they went along. There was little funding for the new discipline, and researchers had no ready outlet for publishing their studies, making it difficult to communicate among themselves about their findings and solutions to methodological problems. Most journal editors were reluctant to publish studies using an experimental method unfamiliar to mainstream social science, which relied on large numbers of subjects and tests of statistical inference.

Despite these problems it was an exciting time, and major new discoveries were being made regularly. For example, many pioneering applications of behavior principles to education occurred during this period (see, e.g., O'Leary & O'Leary, 1972; Ulrich, Stachnik, & Mabry 1974), from which were derived teaching procedures such as contingent teacher praise and attention (Hall, Lund, & Jackson, 1968), token reinforcement systems (Birnbrauer, Wolf, Kidder, & Tague, 1965), curriculum design (Becker, Engelmann, & Thomas, 1975), and programmed instruction (Bijou, Birnbrauer, Kidder, & Tague, 1966; Markle, 1962). The basic methods for reliably improving student performance developed by those early applied behavior analysts provided the foundation for behavioral approaches to curriculum design, instructional methods, classroom management, and the generalization and maintenance of learning that continue to be used decades later (cf., Twyman, 2013).

University programs in behavior analysis were begun in the 1960s and 1970s at Arizona State University, Florida State University, the State University of New York at Stony Brook,

BOX 1.1

What Is Behaviorism?

Don Baer loved the science of behavior. He loved to write about it, and he loved to talk about it. Don was famous for his unparalleled ability to speak extemporaneously about complex philosophical, experimental, and professional issues in a way that always made thorough conceptual, practical, and human sense. He did so with the vocabulary and syntax of a great author and the accomplished delivery of a master storyteller. The only thing Don knew better than his audience was his science.

On three occasions, in three different decades, graduate students and faculty in the special education program at The Ohio State University were fortunate to have Professor Baer serve as Distinguished Guest Faculty for a doctoral seminar, Contemporary Issues in Special Education and Applied Behavior Analysis. The questions and responses that follow were selected from transcripts of two of Professor Baer's three OSU teleconference seminars.

If a person on the street approached you and asked, "What's behaviorism?" how would you reply?

The key point of behaviorism is that what people do can be understood. Traditionally, both the layperson and the psychologist have tried to understand behavior by seeing it as the outcome of what we think, what we feel, what we want, what we calculate, and etcetera. But we don't have to think about behavior that way. We could look upon it as a process that occurs in its own right and has its own causes. And those causes are, very often, found in the external environment.

Behavior analysis is a science of studying how we can arrange our environments so they make very likely the behaviors we want to be probable enough, and they make unlikely the behaviors we want to be improbable. Behaviorism is understanding how the environment works so that we can make ourselves smarter, more organized, more responsible; so we can encounter fewer punishments and fewer disappointments. A central point of behaviorism is this: We can remake our environment to accomplish some of that much more easily than we can remake our inner selves.

An interviewer once asked Edward Teller, the physicist who helped develop the first atomic bomb, "Can you explain to a nonscientist what you find so fascinating about science, particularly physics?" Teller replied, "No." I sense that Teller was suggesting that a nonscientist would not be able to comprehend, understand, or appreciate physics and his fascination with it. If a nonscientist asked you, "What do you find so fascinating about science, particularly the science of human behavior?" what would you say?

Ed Morris organized a symposium on just this topic a couple of years ago at the Association for Behavior Analysis annual convention, and in that symposium, Jack Michael commented on the fact that although one of our discipline's big problems and challenges is communicating with our society about who we are, what we do, and what we can do, he didn't find it reasonable to try to summarize what behavior analysis is to an ordinary person in just a few words. He gave us this example: Imagine a quantum physicist is approached at a cocktail party by someone who asks, "What is quantum physics?" Jack said that the physicist might very well answer, and probably should answer, "I can't tell you in a few words. You should register for my course."

I'm very sympathetic with Jack's argument. But I also know, as someone who's confronted with the politics of relating our discipline to society, that although it may be a true answer, it's not a good answer. It's not an answer that people will hear with any pleasure, or indeed, even accept. . . . Therefore, I think we have to engage in a bit of honest show business. So, if I had to somehow state some connotations of what holds me in the field, I guess I would say that since I was a child I always found my biggest reinforcer was something called understanding. I liked to know how things worked. And of all of the things in the world there are to understand, it became clear to me that the most fascinating was what people do. I started with the usual physical science stuff, and it was intriguing to me to understand how radios work, and how electricity works, and how clocks work, etcetera. But when it became clear to me that we could also learn how people work—not just biologically, but behaviorally—I thought that's the best of all. Surely, everyone must agree that that's the most fascinating subject matter. That there could be a science of behavior, of what we do, of who we are? How could you resist that?

Adapted from "Thursday Afternoons with Don: Selections from Three Teleconference Seminars on Applied Behavior Analysis" by W. L. Heward & C. L. Wood (2003). In K. S. Budd & T. Stokes (Eds.), *A Small Matter of Proof: The Legacy of Donald M. Baer* (pp. 293–310). Reno, NV: Context Press. Used by permission.

the University of Illinois, Indiana University, the University of Kansas, The Ohio State University, the University of Oregon, the University of Southern Illinois, the University of Washington, West Virginia University, and Western Michigan University, among others. Through their teaching and research, faculty at each of these programs made major contributions to the rapid growth of the field.[12]

Two significant events in 1968 mark that year as the formal beginning of contemporary applied behavior analysis. First, the *Journal of Applied Behavior Analysis (JABA)* began publication. *JABA* was the first journal in the United States to deal with applied problems that gave researchers using methodology from the experimental analysis of behavior an outlet for publishing their findings. *JABA* was and continues to be the flagship journal

of applied behavior analysis. Many of the early articles in *JABA* became model demonstrations of how to conduct and interpret applied behavior analysis, which in turn led to improved applications and experimental methodology.

The second major event of 1968 was the publication of the paper "Some Current Dimensions of Applied Behavior Analysis" by Donald M. Baer, Montrose M. Wolf, and Todd R. Risley. These authors, the founding fathers of the new discipline, recommended criteria for judging the adequacy of research and practice in applied behavior analysis and outlined the scope of work they envisioned for those engaged in the science. Their iconic paper is the most widely cited publication in applied behavior analysis and generally regarded as the standard description of the discipline.

CHARACTERISTICS OF APPLIED BEHAVIOR ANALYSIS

Baer, Wolf, and Risley (1968) recommended that applied behavior analysis be *applied, behavioral, analytic, technological, conceptually systematic, effective,* and capable of appropriately *generalized outcomes.* In 1987 Baer and colleagues reported that the "seven self-conscious guides to behavior analytic conduct" (p. 319) they had offered 20 years earlier "remain functional; they still connote the current dimensions of the work usually called applied behavior analysis" (p. 314). The seven dimensions they posed continue to serve as useful and relevant signposts for identifying research in applied behavior analysis.

Applied

The *applied* in applied behavior analysis signals ABA's commitment to effecting improvements in behaviors that enhance and improve people's lives. To meet this criterion, the researcher or practitioner must select behaviors to change that are socially significant for participants: social, language, academic, daily living, self-care, vocational, and/or recreation and leisure behaviors that improve the day-to-day life experience of the participants and/or affect their significant others (parents, teachers, peers, employers) in such a way that they behave more positively with and toward the participant.

Behavioral

At first it may seem superfluous to include such an obvious criterion—of course applied *behavior* analysis must be *behavioral.* However, Baer and colleagues (1968) made three important points relative to the behavioral criterion. First, not just any behavior will do; the behavior chosen for study must be *the* behavior in need of improvement, not a similar behavior that serves as a proxy for the behavior of interest or the subject's verbal description of the behavior. Behavior analysts conduct studies *of* behavior, not studies *about* behavior. For example, in a study evaluating the effects of a program to teach school children to get along with one another, an applied behavior analyst would directly observe and measure clearly defined classes of interactions between and among the children instead of using indirect measures such as the children's answers on a sociogram or responses to a questionnaire about how they believe they get along with one another.

Second, the behavior must be measurable; the precise and reliable measurement of behavior is just as critical in applied research as it is in basic research. Applied researchers must meet the challenge of measuring socially significant behaviors in their natural settings, and they must do so without resorting to the measurement of nonbehavioral substitutes.

Third, when changes in behavior are observed during an investigation, it is necessary to ask whose behavior has changed. Perhaps only the behavior of the observers has changed. "Explicit measurement of the reliability of human observers thus becomes not merely good technique, but a prime criterion of whether the study was appropriately behavioral" (Baer et al., 1968, p. 93). Or perhaps the experimenter's behavior has changed in an unplanned way, making it inappropriate to attribute any observed change in the subject's behavior to the independent variables that were manipulated. The applied behavior analyst should attempt to monitor the behavior of all persons involved in a study.

Analytic

A study in applied behavior analysis is *analytic* when the experimenter has demonstrated a functional relation between the manipulated events and a reliable change in some measurable dimension of the targeted behavior. In other words, the experimenter must be able to control the occurrence and nonoccurrence of the behavior. Sometimes, however, society does not allow the repeated manipulation of important behaviors to satisfy the requirements of experimental method. Therefore, applied behavior analysts must demonstrate control to the greatest extent possible, given the restraints of the setting and behavior; and then they must present the results for judgment by the consumers of the research. The ultimate issue is believability: Has the researcher achieved experimental control to demonstrate a reliable functional relation?

The analytic dimension enables ABA not only to demonstrate effectiveness but also to provide the "acid test proof" of functional and replicable relations between the interventions it recommends and socially significant outcomes.

> Because we are a data- and design-based discipline, we are in the remarkable position of being able to prove that behavior can work in the way that our technology prescribes. We are not theorizing about how behavior *can* work; we are describing systematically how it *has* worked many times in real-world applications, in designs too competent and with measurement systems too reliable and valid to doubt. Our ability to prove that behavior can work that way does not, of course, establish that behavior *cannot* work any other way: we are not in a discipline that can deny any other approaches, only in one that can affirm itself as knowing many of its *sufficient* conditions at the level of experimental proof . . . our subject matter is behavior change, and we can specify some *actionable* sufficient conditions for it. (D. M. Baer, personal communication, October 21, 1982, emphasis in original)

Technological

A study in applied behavior analysis is *technological* when all of its operative procedures are identified and described with sufficient detail and clarity "such that a reader has a fair chance of replicating the application with the same results" (Baer, Blount, Detrich, & Stokes, 1987, p. 320).

> It is not enough to say what is to be done when the subject makes response R_1; it is essential also whenever possible to say what is to be done if the subject makes the alternative responses, R_2, R_3, etc. For example, one may read that temper tantrums in children are often extinguished by closing the child in his room for the duration of the tantrums plus ten minutes. Unless that procedure description also states what should be done if the child tries to leave the room early, or kicks out the window, or smears feces on the walls, or begins to make strangling sounds, etc., it is not precise technological description. (Baer et al., 1968, pp. 95–96)

No matter how powerful its effects in any given study, a behavior change method will be of little value if practitioners are unable to replicate it. The development of a replicable technology of behavior change has been a defining characteristic and continuing goal of ABA from its inception. Behavioral tactics are replicable and teachable to others. Interventions that cannot be replicated with sufficient fidelity to achieve comparable outcomes are not considered part of the technology.

A good check of the technological adequacy of a procedural description is to have a person trained in applied behavior analysis carefully read the description and then act out the procedure in detail. If the person makes any mistakes, adds any operations, omits any steps, or has to ask any questions to clarify the written description, then the description is not sufficiently technological and requires improvement.

Conceptually Systematic

Although Baer and colleagues (1968) did not state so explicitly, a defining characteristic of applied behavior analysis concerns the types of interventions used to improve behavior. Although an infinite number of tactics and specific procedures can be used to alter behavior, almost all are derivatives and/or combinations of a relatively few basic principles of behavior. Thus, Baer and colleagues recommended that research reports of applied behavior analysis be *conceptually systematic,* meaning that the procedures for changing behavior and any interpretations of how or why those procedures were effective should be described in terms of the relevant principle(s) from which they were derived.

Baer and colleagues (1968) provided a strong rationale for the use of conceptual systems in applied behavior analysis. First, relating specific procedures to basic principles might enable the research consumer to derive other similar procedures from the same principle(s). Second, conceptual systems are needed if a technology is to become an integrated discipline instead of a "collection of tricks." Loosely related collections of tricks do not lend themselves to systematic expansion, and they are difficult to learn and to teach.

Effective

An effective application of behavioral techniques must improve the behavior under investigation to a practical degree. "In application, the theoretical importance of a variable is usually not at issue. Its practical importance, specifically its power in altering behavior enough to be socially important, is the essential criterion" (Baer et al., 1968, p. 96). Whereas some investigations produce results of theoretical importance or statistical significance, to be judged *effective* an applied behavior analysis study must produce behavior changes that reach clinical or social significance.

How much a given behavior of a given subject needs to change for the improvement to be considered socially important is a practical question. Baer and colleagues stated that the answer is most likely to come from the people who must deal with the behavior; they should be asked how much the behavior needs to change. The necessity of producing behavioral changes that are meaningful to the participant and/or those in the participant's environment has pushed behavior analysts to search for "robust" variables, interventions that produce large and consistent effects on behavior (Baer, 1977a).

When they revisited the dimension of effectiveness 20 years later, Baer and colleagues (1987) recommended that the effectiveness of ABA also be judged by a second kind of outcome: the extent to which changes in the target behaviors result in noticeable changes in the reasons those behaviors were selected for change originally. If such changes in the subjects' lives do not occur, ABA may achieve one level of effectiveness yet fail to achieve a critical form of social validity (Wolf, 1978).

> We may have taught many social skills without examining whether they actually furthered the subject's social life; many courtesy skills without examining whether anyone actually noticed or cared; many safety skills without examining whether the subject was actually safer thereafter; many language skills without measuring whether the subject actually used them to interact differently than before; many on-task skills without measuring the actual value of those tasks; and, in general, many survival skills without examining the subject's actual subsequent survival. (Baer et al., 1987, p. 322)

Generality

A behavior change has *generality* if it lasts over time, appears in environments other than the one in which the intervention that initially produced it was implemented, and/or spreads to other behaviors not directly treated by the intervention. A behavior change that continues after the original treatment procedures are withdrawn has generality. And generality is

evident when changes in targeted behavior occur in nontreatment settings or situations as a function of treatment procedures. Generality also exists when behaviors change that were not the focus of the intervention. Although not all instances of generality are adaptive (e.g., a beginning reader who has just learned to make the sound for the letter *p* in words such as *pet* and *ripe,* might make the same sound when seeing the letter *p* in the word *phone*), desirable generalized behavior changes are important outcomes of an applied behavior analysis program because they represent additional dividends in terms of behavioral improvement. Strategies and tactics for promoting desirable generalization of behavior changes are detailed in Chapter 30.

More than 50 years have passed since Baer, Wolf, and Risley (1968) proposed these seven dimensions as the defining features of applied behavior analysis. For insightful discussion of the usefulness of these dimensions for contemporary ABA research, see Axelrod (2017), Cataldo (2017), Critchfield and Reed (2017), and Friman (2017).

Other Characteristics of ABA

Applied behavior analysis offers society an approach toward solving problems that is accountable, public, doable, empowering, and optimistic (Heward, 2005). These characteristics should make behavior analysts "feel good" and increase the extent to which decision makers and consumers in many areas look to applied behavior analysis as a valuable and important source of knowledge for achieving improved outcomes.

Accountable

Applied behavior analysts' commitment to effectiveness, their focus on accessible environmental variables that reliably influence behavior, and their reliance on direct and frequent measurement to detect changes in behavior yield an inescapable and socially valuable form of accountability. Direct and frequent measurement—the foundation and most important component of ABA practices—enables behavior analysts to identify their successes and, equally important, their failures so they can make changes in an effort to change failure to success (Bushell & Baer, 1994; Greenwood & Maheady, 1997).

> Failure is always informative in the logic of behavior analysis, just as it is in engineering. The constant reaction to lack of progress [is] a definitive hallmark of ABA. (Baer, 2005, p. 8)

Gambrill (2003) described the sense of accountability and self-correcting nature of applied behavior analysis very well.

> Applied behavior analysis is a scientific approach to understanding behavior in which we guess and critically test ideas, rather than guess and guess again. It is a process for solving problems in which we learn from our mistakes. Here, false knowledge and inert knowledge are not valued. (p. 67)

Public

"Everything about ABA is visible and public, explicit and straightforward. . . . ABA entails no ephemeral, mystical, or metaphysical explanations; there are no hidden treatments; there is no magic" (Heward, 2005, p. 322). The transparent, public nature of ABA should raise its value in fields such as education, parenting and child care, employee productivity, geriatrics, health and safety, and social work—to name only a few—whose goals, methods, and outcomes are of vital interest to many constituencies.

Doable

Classroom teachers, parents, coaches, workplace supervisors, and sometimes the participants themselves implemented the interventions found effective in many ABA studies. This demonstrates the pragmatic element of ABA. "Although 'doing ABA' requires far more than learning to administer a few simple procedures, it is not prohibitively complicated or arduous. As many teachers have noted, implementing behavioral strategies in the classroom . . . might best be described as good old-fashioned hard work" (Heward, 2005, p. 322).

Empowering

ABA gives practitioners real tools that work. Knowing how to do something and having the tools to accomplish it instills confidence in practitioners. Seeing the data showing behavioral improvements in one's clients, students, or teammates, or in oneself, not only feels good but also raises one's self-assurance in assuming even more difficult challenges in the future.

Optimistic

Practitioners knowledgeable and skilled in behavior analysis have genuine cause to be optimistic for four reasons. First, as Strain and Joseph (2004) noted:

> The environmental view promoted by behaviorism is essentially optimistic; it suggests that (except for gross genetic factors) all individuals possess roughly equal potential. Rather than assuming that individuals have some essential internal characteristic, behaviorists assume that poor outcomes originate in the way the environment and experience shaped the individual's current behavior. Once these environmental and experiential factors are identified, we can design prevention and intervention programs to improve the outcomes. . . . Thus, the emphasis on external control in the behavioral approach . . . offers a conceptual model that celebrates the possibilities for each individual. (Strain et al., 1992, p. 58)

Second, direct and continuous measurement enables practitioners to detect small improvements in performance that might otherwise be overlooked. Third, the more often a practitioner uses behavioral tactics with positive outcomes (the most common result of behaviorally based interventions), the more optimistic she becomes about the prospects for future success.

> A sense of optimism, expressed by the question "Why not?" has been a central part of ABA and has had an enormous impact on its development from its earliest days. Why can't we teach a person who does not yet talk to talk? Why shouldn't we go ahead and try to change the environments of young children so that they will display more creativity? Why would we assume that this person with a developmental disability could not learn to do the same things that many of us do? Why not try to do it? (Heward, 2005, p. 323)

Fourth, ABA's peer-reviewed literature provides many examples of success in teaching students who had been considered unteachable. ABA's continuous record of achievements evokes a legitimate feeling of optimism that future developments will yield solutions to behavioral challenges that are currently beyond the existing technology. For example, in response to the perspective that some people have disabilities so severe and profound that they should be viewed as ineducable, Don Baer offered this perspective:

> Some of us have ignored both the thesis that all persons are educable and the thesis that some persons are ineducable, and instead have experimented with ways to teach some previously unteachable people. Those experiments have steadily reduced the size of the apparently ineducable group relative to the obviously educable group. Clearly, we have not finished that adventure. Why predict its outcome, when we could simply pursue it, and just as well without a prediction? Why not pursue it to see if there comes a day when there is such a small class of apparently ineducable persons left that it consists of one elderly person who is put forward as ineducable. If that day comes, it will be a very nice day. And the next day will be even better. (D. M. Baer, February 15, 2002, personal communication, as cited in Heward, Alber-Morgan, & Konrad, 2017, p. 404)

A DEFINITION OF APPLIED BEHAVIOR ANALYSIS

We began this chapter by stating that applied behavior analysis is a science with a dual purpose of understanding and improving socially important behavior. We then described attitudes, assumptions, and methods that are fundamental to scientific inquiry, briefly reviewed the development of the science and philosophy of behavior analysis, and examined the characteristics of ABA. All of that provided necessary context for offering the following definition of applied behavior analysis:

> **Applied behavior analysis** is the science in which tactics derived from the principles of behavior are applied systematically to improve socially significant behavior and experimentation is used to identify the variables responsible for behavior change.

This definition includes six key components. First, applied behavior analysis is a science, which means ABA researchers and practitioners are guided by the attitudes and methods of scientific inquiry. Second, all behavior change procedures are described and implemented in a systematic, technological manner. Third, not any means of changing behavior qualifies as applied behavior analysis: The field circumscribes only those tactics conceptually derived from the basic principles of behavior. Fourth, the focus of applied behavior analysis is socially significant behavior. The fifth and sixth parts of the definition specify the twin goals of applied behavior analysis: improvement and understanding. Applied behavior analysis seeks to make meaningful improvement in important behavior and to achieve an analysis of the factors responsible for that improvement.[13]

Four Interrelated Domains of Behavior Analytic Science and Professional Practice Guided by That Science

The science of behavior analysis and its application to human problems consists of four domains: the three branches of behavior analysis—radical behaviorism, EAB, and ABA—and professional practice in various fields that is informed and guided by that science. Figure 1.2 identifies defining features and characteristics of each domain. Although most behavior analysts work primarily in one or two of the domains shown in Figure 1.2, it is common for a behavior analyst to function in multiple domains at one time or another (Hawkins & Anderson, 2002; Moore & Cooper, 2003).

Radical behaviorism, the philosophical domain of behavior analysis, entails pursuing theoretical and conceptual issues. Examples of theoretical and conceptual behavior analysis are Glenn's (2004) exploration on the interplay of behavioral and cultural practices; Schlinger's (2008b) analysis of listening as behaving verbally; Dillenburger and Keenan's (2005) discussion of bereavement; Layng's (2017) theoretical account of emotions and emotional behavior; and M. Malott's (2016) examination of leadership.

The experimental analysis of behavior is the basic research branch of the science. Basic research consists of experiments in (mostly) laboratory settings with both human and nonhuman subjects with a goal of discovering, extending, and clarifying fundamental principles of behavior. Examples of the wide range of topics investigated by EAB researchers include how organisms direct their attention in complex environments (Shahan, 2013); choice making (Mazur & Fantino, 2014); remembering and forgetting (White, 2013); delay discounting (Odum, 2011); and variability of operant behavior (Neuringer & Jensen, 2013).[14, 15]

Applied behavior analysts conduct experiments aimed at discovering and clarifying functional relations between socially significant behavior and its controlling variables, with which they can contribute to the further development of humane and effective technologies of behavior change. Examples of ABA research include Tarbox, Wallace, and Williams's (2003)

Figure 1.2 Comparisons and relationships among the four domains of behavior analysis science and practice.

	Radical Behaviorism	**Experimental Analysis of Behavior (EAB)**	*Translational Behavior Analysis**	**Applied Behavior Analysis (ABA)**	**Practice Guided by Behavior Analysis**
Province	Theory and philosophy	Basic research		Applied research	Helping people behave more successfully
Primary activity	Conceptual and philosophical analysis	Design, conduct, interpret, and report basic experiments		Design, conduct, interpret, and report applied experiments	Design, implement, and evaluate behavior change programs
Primary goal and product	Theoretical account of all behavior consistent with existing data	Discover and clarify basic principles of behavior; functional relations between behavior and controlling variables		A technology for improving socially significant behavior; functional relations between socially significant behavior and controlling variables	Improvements in the lives of participants/clients as a result of changes in their behavior
Secondary goals	Identify areas in which empirical data are absent and/or conflict and suggest resolutions	Identify questions for EAB and/or ABA to investigate further; raise theoretical issues		Identify questions for EAB and/or ABA to investigate further; raise theoretical issues	Increased efficiency in achieving primary goal; may identify questions for ABA and EAB
Agreement with existing database	As much as possible, but theory must go beyond database by design	Complete—Although differences among data sets exist, EAB provides the basic research database		Complete—Although differences among data sets exist, ABA provides the applied research database	As much as possible, but practitioners must often deal with situations not covered by existing data
Testability	Partially—All behavior and variables of interest are not accessible (e.g., phylogenic contingencies)	Mostly—Technical limitations preclude measurement and experimental manipulation of some variables		Mostly—Same limitations as EAB plus those posed by applied settings (e.g., ethical concerns, uncontrolled events)	Partially—All behavior and variables of interest are not accessible (e.g., a student's home life)
Scope	**Most** ←	←		→	→ **Least**
	Wide scope because theory attempts to account for all behavior	As much scope as the EAB database enables		As much scope as the ABA database enables	Narrow scope because practitioner's primary focus is helping the specific situation
Precision	**Least** ←	←		→	→ **Most**
	Minimal precision is possible because experimental data do not exist for all behavior encompassed by theory	As much precision as EAB's current technology for experimental control and the researcher's skills enable		As much precision as ABA's current technology for experimental control and the researcher's skills enable	Maximum precision is needed to change behavior most effectively in specific instance

*Translational behavior analysis bridges between EAB and ABA (see McIlvane, 2009).

assessment and treatment of elopement (running or walking away from a caregiver without permission) by individuals with disabilities; Romanowich and Lamb's (2015) research on the effects of various schedules of reinforcement on smokers' abstinence; Crabtree, Alber-Morgan, and Konrad's (2010) experiment on the effects of self-monitoring of story elements on the reading comprehension of high school students with learning disabilities; and Poling and colleagues' teaching giant African pouched rats to find unexploded landmines (Poling et al., 2010), detect patients infected with tuberculosis (Poling et al., 2017), and find people trapped under debris from collapsed structures resulting from natural disasters, acts of war or terrorism, or engineering mistakes (La Londe et al., 2015).

"Behavior analysis is not merely the sum of its basic and applied research and conceptual programs. It is their interrelationship, wherein each branch draws strength and integrity from the others. With the unity of behavior analysis clarified, the whole of behavior analysis emerges as greater than the sum of its parts" (Morris, Todd, Midgley, Schneider, & Johnson, 1990, p. 136).

Translational research is evidence of the symbiotic relation between the basic and applied domains. Translational research includes basic research that "bridges" basic and applied areas and applied research that translates the knowledge derived from basic research "into state-of-the-art clinical practices for use in the community" (Lerman, 2003, p. 415). In a review of translational research on token reinforcement, Hackenberg (2018) suggested that all three domains of behavior analysis science benefit when basic and applied research inform each other.

> Unlike standard translational research, based on a unidirectional model in which the analysis moves from laboratory to the applied realm, work in the area of token systems is best served by a bidirectional interplay between laboratory and applied research, where applied questions inspire research on basic mechanisms. When based on and contributing to an analysis, applied research on token economies can be on the leading edge of theoretical advances, helping set the scientific research agenda. (p. 393)[16]

The delivery of professional services informed by behavior analysis occurs in the fourth domain. Behavior analytic practitioners design, implement, and evaluate behavior change programs that consist of behavior change tactics derived from fundamental principles of behavior discovered by basic researchers and experimentally validated for their effects on socially significant behavior by applied researchers. An example is when a therapist providing home-based treatment for a child with autism embeds frequent opportunities for the child to use emerging social and language skills in the context of naturalistic, daily routines and ensures that the child's responses are followed with reinforcing events. Another example is a classroom teacher trained in behavior analysis who uses positive reinforcement and stimulus fading to teach students to identify and classify fish into their respective species by the shape, size, and location of their fins.

Although each of the four domains can be defined and practiced in its own right, none of the domains are, or should be, completely independent of and uninformed by developments in the others. Both the science and the effective application of its findings benefit when the four domains are interrelated and influence one another.

The Promise and Potential of ABA

In a paper titled "A Futurist Perspective for Applied Behavior Analysis," Jon Bailey (2000) stated that

> It seems to me that applied behavior analysis is more relevant than ever before and that it offers our citizens, parents, teachers, and corporate and government leaders advantages that cannot be matched by any other psychological approach. . . . I know of no other approach in psychology that can boast state-of-the-art solutions to the most troubling social ills of the day. (p. 477)

We, too, believe that ABA's pragmatic, natural science approach to discovering environmental variables that reliably influence socially significant behavior and to developing a technology to take practical advantage of those discoveries offers humankind its best hope for solving many of its problems. It is important to recognize that the knowledge afforded by ABA of "how behavior works" is incomplete, even at the level of fundamental principles, as is the technology for changing behavior derived from that knowledge. There are aspects of behavior about which relatively little is known, and additional research, both basic and applied, is needed to clarify, extend, and fine-tune all existing knowledge.

Nevertheless, applied behavior analysis research and practice have improved human performance and the quality of participants' lives across a wide range of areas. Figure 1.3 lists some of the more than 230 applied behavior analysis research topics identified by an informal literature search and survey (Heward & Critchfield, 2019). As diverse and impressive as this A-to-Z list of topics may seem, no behavioral problem or learning goal has been solved completely, and many important problems and challenges await analysis. The remainder of this text presents a foundation of knowledge that can lead to a fuller and more complete understanding of the still young and promising science of applied behavior analysis.

Figure 1.3 ABA: Improving people's quality of life from A to Z.

A
ADHD (Bicard & Neef, 2002), aging (Baker & LeBlanc, 2014), aggression (Brosnan & Healy, 2011), AIDS (DeVries et al., 1991), alcohol abuse (Fournier et al., 2004), Alzheimer's (LeBlanc et al., 2006), anorexia nervosa (Solanto et al., 1994), animal training (Protopopova et al., 2016), autism (Ahearn & Tiger, 2013), aviation safety (Rantz & Van Houten, 2013)

B
baseball (Heward, 1978), basketball (Kladopoulos & McComas, 2001), bedtime refusal (Friman et al., 1999), behavioral pharmacology (Roll, 2014), bicycle safety (Okinaka & Shimazaki, 2011), blackjack skills (Speelman et al., 2015), braille (Scheithauer & Tiger, 2014), breast cancer detection (Bones et al., 2016), bruxism (Barnoy et al., 2009), bullying (Ross et al., 2009)

C
caffeinism (Foxx & Rubinoff, 1979), cancer prevention (Lombard et al., 1991), child abuse (Van Camp et al., 2003), climate change (Heward & Chance, 2010), coaching athletes (Stokes et al., 2010), college teaching (Kellum et al., 2001), compassion (Geller, 2012), cooperative learning (Maheady et al., 2006), creativity (Winston & Baker, 1985), crying (Bowman, 2013), culturally responsive social skills instruction (Lo et al., 2015)

D
dance (Quinn et al., 2015), delusional speech (Travis & Sturmey, 2010), dementia (Engelman et al., 1999), depression (Follette & Darrow, 2014), developmental disabilities (Kurtz & Lind, 2013), diabetes (Raiff et al., 2016), drug addiction (Silverman et al., 2011), dyslexia (Denton & Meindl, 2016)

E
education (Heward et al., 2005), elopement (Kodak et al., 2004), energy conservation (Staats et al., 2000), enuresis (Friman & Jones, 2005)

F
fear of dental procedures (Conyers et al., 2004), feeding disorders (Volkert & Piazza, 2012), figure skating (Hume et al., 1985), fire safety (Garcia et al., 2016), food bank donations (Farrimond & Leland, 2006), football (Ward & Carnes, 2002), foster-family care (Hawkins et al., 1985)

G
gambling (Dixon et al., 2015), gamification (Morford et al., 2014) gender-based violence (Szabo et al., 2019), gerontology (Gallagher & Keenan, 2000), golf (Simek et al., 1994), gun safety (Miltenberger et al., 2005)

H
hairpulling (Rapp et al., 1999), handwriting (Trap et al., 1978), happiness (Parsons et al., 2012), headaches (Fitterling et al., 1988), highway safety (Van Houten et al., 1985), homework (Alber et al., 2002), horseback riding (Kelley & Miltenberger, 2016), horse training (Fox & Belding, 2015), hostage negotiation (Hughes, 2006), hygiene (Fournier & Berry, 2013)

I
impulsivity (Barry & Messer, 2003), incontinence (Adkins & Mathews, 1997), industrial safety (Fox et al., 1987), infant care (Dachman, et al., 1986), infection control (Babcock et al., 1992), intellectual disabilities (Frederick et al., 2013)

J
job finding (Azrin et al., 1975), joint attention (Taylor & Hoch, 2008), juvenile justice (Kirigin et al., 1982)

K
keyboard training (DeFulio, 2011), kidnapping prevention (Gunby et al., 2010), kitchen skills (Trask-Tyler et al., 1994), kleptomania (Kohn, 2006)

L
landmine detection (Edwards et al., 2015), language acquisition (Drasgow, et al., 1998), learning disabilities (Wolfe et al., 2000), leisure skills Schleien et al., 1981), littering (Powers et al., 1973), lockdown drill procedures (Dickson & Vargo, 2017)

M
martial arts (BenitezSantiago & Miltenberger, 2016), math (Hunter et al., 2016), medical procedures (Hagopian & Thompson, 1999), medical training (Levy et al., 2016), mental health (A-tjak et al., 2015), music skills (Griffith et al., 2018), myopia (Collins et al., 1981)

N
nail biting (Heffernan & Lyons, 2016), nervous habits (see nail biting, hair pulling), noise (Ring et al., 2014), noncompliance (Mace et al., 1988), nutrition (Horne et al., 2009)

O
obesity (De Luca & Holborn, 1992), observational learning (DeQuinzio & Taylor, 2015), obsessive-compulsive disorder (Penney et al., 2016), organizational behavior management (Rodriguez, 2011), otitis media (O'Reilly, 1997)

P
panic disorders, parenting (Miltenberger & Crosland, 2014), pediatrics (Friman & Piazza, 2011), phobias (Tyner et al., 2016), physical activity (Kuhl et al., 2015), physical education (McKenzie et al., 2009), pica (Hagopian et al., 2011), play skills (Davis-Temple et al., 2014), Prader-Willi syndrome (Page et al., 1983), problem solving (Axe et al., in press), procrastination (Johnson et al., 2016), public health (Biglan & Glenn, 2013), public speaking (Mancuso & Miltenberger, 2016)

Q
quality control (Kortick & O'Brien, 1996) quantitative analysis skills (Fienup & Critchfield, 2010), question asking/answering (Ingvarsson et al., 2007)

R
x-ray sheilding, reading (Twyman et al., 2005), recycling (O'Conner et al., 2010), robbery (Schnelle et al., 1979), rugby (Mellalieu et al., 2006), rumination (Woods et al., 2013)

S
safe sex (Honnen & Kleinke, 1990), school-wide behavior support (Freeman et al., 2016), seat-belt use (Van Houten et al., 2010), second language acquisition (May et al., 2016), self-injury (Lerman & Iwata, 1993), self-management (Reynolds et al., 2014), sexual abuse (Lumley et al., 1998), sharing (Marzullo-Kerth et al., 2011), sleep disorders (Piazza et al., 1997), soccer (Brobst & Ward, 2002), spelling (McNeish et al., 1992), stereotypy (Ahearn et al., 2003), stuttering (Wagaman et al., 1995), substance abuse (Roll et al., 2009), sustainability (Leeming et al., 2013), swimming (Hume & Crossman, 1992)

T
tantrums (Williams, 1959), TB detection (Poling et al., 2017), teacher training (Kretlow et al., 2012), thumb sucking (Friman, 2000), tobacco use (Romanowich & Lamb, 2015), toileting (Greer et al., 2016), toothbrushing (Poche et al., 1982), Tourette syndrome (Azrin & Peterson, 1988), track and field (Scott et al., 1997), traumatic brain injury (Heinicke et al., 2009)

U
urban search and rescue (Edwards et al., 2016), universal precautions (Luke & Alavosius, 2011)

V
vandalism (Mayer et al., 1983), visual acuity (Collins et al., 1981), vocal tics (Wagaman et al., 1995), vocational training (Cullen et al., 2017), voice disorders, (Shriberg, 1971)"

W
weight loss (VanWormer, 2004), workplace safety (Abernathy & Lattal, 2014), writer's block (Didden et al., 2007), writing skills (Hansen & Wills, 2014)

X
x-ray shielding (Greene & Neistat, 1983)

Y
yielding to pedestrians (Bennett et al., 2014), yoga (Downs, 2015), youth sports (Luiselli et al., 2011)

Z
zoo animal welfare (Maple & Segura, 2015)

Notes: Adapted from "ABA from A-to-Z" by W. L. Heward and T. S. Critchfield, 2019. Manuscript in progress. References with three or more coauthors cited as "et al."

SUMMARY

Science: Basic Characteristics and a Definition

1. Different types of scientific investigations yield knowledge that enables the description, prediction, and/or control of the phenomena studied.
2. Descriptive studies yield a collection of facts about the observed events that can be quantified, classified, and examined for possible relations with other known facts.
3. Knowledge gained from a study that finds the systematic covariation between two events—termed a correlation—can be used to predict the probability that one event will occur based on the occurrence of the other event.
4. Results of experiments that show that specific manipulations of one event (the independent variable) produce a reliable change in another event (the dependent variable), and that the change in the dependent variable was unlikely the result of extraneous factors (confounding variables)—a finding known as a functional relation—can be used to control the phenomena under investigation.
5. The behavior of scientists in all fields is characterized by a common set of assumptions and attitudes:
   - Determinism—the assumption that the universe is a lawful and orderly place in which phenomena occur as a result of other events.
   - Empiricism—the objective observation of the phenomena of interest.
   - Experimentation—the controlled comparison of some measure of the phenomenon of interest (the dependent variable) under two or more different conditions in which only one factor at a time (the independent variable) differs from one condition to another.
   - Replication—repeating experiments (and independent variable conditions within experiments) to determine the reliability and usefulness of findings.
   - Parsimony—simple, logical explanations must be ruled out, experimentally or conceptually, before more complex or abstract explanations are considered.
   - Philosophic doubt—continually questioning the truthfulness and validity of all scientific theory and knowledge.

A Brief History of Behavior Analysis

6. Behavior analysis consists of three major branches: behaviorism, the experimental analysis of behavior (EAB), and applied behavior analysis (ABA).
7. Watson espoused an early form of behaviorism known as stimulus–response (S–R) psychology, which did not account for behavior without obvious antecedent causes.
8. Skinner founded the experimental analysis of behavior (EAB), a natural science approach for discovering orderly and reliable relations between behavior and the environmental variables of which it is a function.
9. EAB is characterized by these methodological features:
   - Rate of response is the most common dependent variable.
   - Repeated or continuous measurement is made of carefully defined response classes.
   - Within-subject experimental comparisons are used instead of designs comparing the behavior of experimental and control groups.
   - The visual analysis of graphed data is preferred over statistical inference.
   - A description of functional relations is valued over formal theory testing.
10. Through thousands of laboratory experiments, Skinner and his colleagues and students discovered and verified the basic principles of operant behavior that provide the empirical foundation for behavior analysis today.
11. Skinner wrote extensively about a philosophy for a science of behavior he called radical behaviorism. Radical behaviorism attempts to explain all behavior, including private events such as thinking and feeling.
12. Methodological behaviorism is a philosophical position that considers behavioral events that cannot be publicly observed to be outside the realm of the science.
13. Mentalism is an approach to understanding behavior that assumes that a mental, or "inner," dimension exists that differs from a behavioral dimension and that phenomena in this dimension either directly cause or at least mediate some forms of behavior; it relies on hypothetical constructs and explanatory fictions.
14. Pragmatism is the philosophical position that the truth or value of a scientific statement is determined by the extent to which it promotes effective action.
15. The first published report of the application of operant conditioning with a human subject was a study by Fuller (1949), in which an arm-raising response was conditioned in an adolescent with profound disabilities.
16. The formal beginnings of applied behavior analysis can be traced to 1959 and the publication of Ayllon and Michael's article "The Psychiatric Nurse as a Behavioral Engineer."
17. Contemporary applied behavior analysis (ABA) began in 1968 with the publication of the first issue of the *Journal of Applied Behavior Analysis (JABA).*

Characteristics of Applied Behavior Analysis

18. Baer, Wolf, and Risley (1968) stated that a research study or behavior change program should meet seven defining dimensions to be considered applied behavior analysis:

- Applied—investigates socially significant behaviors with immediate importance to the subject(s).
- Behavioral—entails precise measurement of the actual behavior in need of improvement and documents that it was the subject's behavior that changed.
- Analytic—demonstrates experimental control over the occurrence and nonoccurrence of the behavior—that is, if a functional relation is demonstrated.
- Technological—the written description of all procedures used in the study is sufficiently complete and detailed to enable others to replicate it.
- Conceptually systematic—behavior change interventions are derived from basic principles of behavior.
- Effective—improves behavior sufficiently to produce practical results for the participant/client.
- Generality—produces behavior changes that last over time, appear in other environments, and/or spread to other behaviors.

19. ABA offers society an approach toward solving many of its problems that is accountable, public, doable, empowering, and optimistic.

A Definition of Applied Behavior Analysis

20. Applied behavior analysis is the science in which tactics derived from the principles of behavior are applied systematically to improve socially significant behavior and experimentation is used to identify the variables responsible for behavior change.

21. Behavior analysts work in one or more of four interrelated domains: behaviorism (theoretical and philosophical issues), the experimental analysis of behavior (basic research), applied behavior analysis (applied research), and professional practice (providing behavior analytic services to consumers).

22. Translational research bridges basic and applied research and informs both domains.

23. ABA's natural science approach to discovering environmental variables that reliably influence socially significant behavior and developing a technology to take practical advantage of those discoveries offers humankind its best hope for solving many of its problems.

24. Applied behavior analysis research and practice have improved human performance and the quality of participants' lives across a wide range of areas, but no problem has been solved completely, and many important problems, challenges, and opportunities remain.

KEY TERMS

applied behavior analysis (ABA)
behaviorism
determinism
empiricism
experiment
experimental analysis of behavior (EAB)
explanatory fiction
functional analysis
functional relation
hypothetical construct
mentalism
methodological behaviorism
parsimony
philosophic doubt
pragmatism
radical behaviorism
replication
science

MULTIPLE-CHOICE QUESTIONS

1. The level of investigation that involves the collection of facts about observed events that can be quantified, classified, and examined for possible relations with other known facts, and often suggests hypotheses or questions for additional research is:

   a. Prediction
   b. Experimentation
   c. Description
   d. Control

   Hint: (See "Science: Basic Characteristics and a Definition")

2. The level of investigation that demonstrates correlation between events and is based on repeated observations is:

   a. Prediction
   b. Experimentation
   c. Description
   d. Control

   Hint: (See "Science: Basic Characteristics and a Definition")

3. The level of investigation in which functional relations can be derived is:
   a. Prediction
   b. Experimentation
   c. Description
   d. Control
   Hint: (See "Science: Basic Characteristics and a Definition")

4. The overall goal of ________ is to achieve a thorough understanding of the phenomenon under study.
   a. Behavior Analysis
   b. Science
   c. Experimentation
   d. Functional relationships
   Hint: (See "Science: Basic Characteristics and a Definition")

5. A functional relation means that:
   a. Specific change in the independent variable can reliably be produced by specific manipulations in the dependent variable, and the change in the independent variable was unlikely to be the result of confounding variables.
   b. Specific change in the dependent variable can reliably be produced by specific manipulations in the independent variable, and the change in the independent variable was unlikely to be the result of confounding variables.
   c. Specific change in the dependent variable can reliably be produced by specific manipulations in the independent variable, and the change in the dependent variable was unlikely to be the result of confounding variables.
   d. Specific change in the dependent variable can reliably be produced by specific manipulations in the confounding variable, and the change in the dependent variable was unlikely to be the result of the independent variable.
   Hint: (See "Science: Basic Characteristics and a Definition")

6. This is the assumption upon which science is predicated, the idea that the universe is a lawful and orderly place in which all phenomena occur as a result of other events.
   a. Mentalism
   b. Determinism
   c. Empiricism
   d. Philosophic Doubt
   Hint: ("Science: Basic Characteristics and a Definition")

7. This is the idea that simple, logical explanations must be ruled out, experimentally or conceptually, before more complex or abstract explanations are considered.
   a. Philosophic Doubt
   b. Experimentation
   c. Replication
   d. Parsimony
   Hint: ("Science: Basic Characteristics and a Definition")

8. This is the branch of behavior analysis that focuses on basic research:
   a. Applied behavior analysis
   b. Behaviorism
   c. Experimental analysis of behavior
   d. Radical behaviorism
   Hint: (See "A Brief History of Behavior Analysis")

9. The S-R-S model of psychology is also known as:
   a. Three-term contingency
   b. Watsonian psychology
   c. Respondent behavior model
   d. Reflexive behavior model
   Hint: (See "A Brief History of Behavior Analysis")

10. This person is considered to be the founder of experimental analysis of behavior.
    a. John B. Watson
    b. B.F. Skinner
    c. Ivan Pavlov
    d. Don Baer
    Hint: (See "A Brief History of Behavior Analysis")

11. This is the approach to understanding behavior that assumes that a mental or "inner" dimension exists that differs from a behavioral dimension and that phenomena in this dimension either directly cause or at least mediate some forms of behavior.
    a. Radical behaviorism
    b. Methodological behaviorism
    c. Structuralism
    d. Mentalism
    Hint: (See "A Brief History of Behavior Analysis")

12. These events marked the formal beginning of contemporary applied behavior analysis.
    a. "Some Current Dimensions of Applied Behavior Analysis," by Baer, Wolf, and Risley, was published and "The Psychiatric Nurse as a Behavioral Engineer," by Ayllon and Michael, was published.
    b. "The Psychiatric Nurse as a Behavioral Engineer," by Ayllon and Michael, was published and Fuller conducted a study in which human application of operant behavior occurred.
    c. The Journal of Applied Behavior Analysis began publication and "Some Current Dimensions of Applied Behavior Analysis," by Baer, Wolf, and Risley, was published.
    d. The Journal of Applied Behavior Analysis began publication and "The Psychiatric Nurse as a Behavioral Engineer," by Ayllon and Michael, was published.
    Hint: (See "A Brief History of Behavior Analysis")

13. This is the defining characteristic of behavior analysis that focuses on investigating socially significant behaviors with immediate importance to the participant(s).
    a. Effective
    b. Applied
    c. Behavioral
    d. Analytical
    Hint: (See "Characteristics of Applied Behavior Analysis")

14. This is the defining characteristic of applied behavior analysis that demonstrates experimental control over the occurrence and non-occurrence of the behavior.
    a. Effective
    b. Analytic
    c. Applied
    d. Experimentation
    Hint: (See "Characteristics of Applied Behavior Analysis")

ESSAY-TYPE QUESTIONS

1. Describe the goal of science and the three different levels of investigation.
   Hint: ("Science: Basic Characteristics and a Definition")
2. State and describe the different scientific attitudes.
   Hint: ("Science: Basic Characteristics and a Definition")
3. Discuss similarities and differences between different perspectives on behavior (i.e., mentalism, radical behaviorism, methodological behaviorism, and structuralism).
   Hint: (See "A Brief History of Behavior Analysis")
4. Discuss the origin of applied behavior analysis, including early history of the field and defining events of contemporary applied behavior analysis.
   Hint: (See "A Brief History of Behavior Analysis")
5. Compare and contrast the different branches of behavior analysis.
   Hint: (See "A Brief History of Behavior Analysis")
6. State and describe each of the defining dimensions of applied behavior analysis.
   Hint: (See "Characteristics of Applied Behavior Analysis")

NOTES

1. Spurious correlations result from measures of two randomly chosen variables closely tracking each other (e.g., rise in annual organic food sales and increased incidence of autism [Redditor Jasonp55, 2018], per capita margarine consumption and the divorce rate in Maine [Vigen, 2015]). For more silly examples, see Vigen (2015); for detailed explanation of the fallacy, see West, Bergstrom, and Bergstrom (2010).
2. Skinner (1953) noted that although telescopes and cyclotrons give us a "dramatic picture of science in action" (p. 12), and science could not have advanced very far without them, such devices and apparatus are not science themselves. "Nor is science to be identified with precise measurement. We can measure and be mathematical without being scientific at all, just as we may be scientific without these aids" (p. 12). Scientific instruments bring scientists into greater contact with their subject matter and, with measurement and mathematics, enable a more precise description and control of key variables.
3. Informative and interesting descriptions of the history of behavior analysis can be found in Goodall (1972); Guericio (2018); Hackenberg (1995); Michael (2004); Morris, Todd, Midgley, Schneider, and Johnson (1990); Mountjoy and Cone (1997); Risley (2005); Sidman (2002); Skinner (1956, 1979); Stokes (2003); Vargas, Vargas, and Knapp (2017); and in a special section of articles in the Fall 2003 issue of *The Behavior Analyst*.
4. For interesting biographies and scholarly examinations of J. B. Watson's life and contributions to behavior analysis, see Catania (1993); Morris (2013); and Morrow (2017).
5. In *The Behavior of Organisms,* Skinner called the conditioning of respondent behavior Type S conditioning and the conditioning of operant behavior Type R conditioning, but these terms were soon dropped. Respondent and operant conditioning and the three-term contingency are further defined and discussed in Chapter 2.
6. Most of the methodological features of the experimental approach pioneered by Skinner (e.g., rate of response as the primary dependent variable, within-subject experimental comparisons, visual analysis of graphic data displays) continue to characterize both basic and applied research in behavior analysis (e.g., Fisher, Piazza, & Roane, 2011; Lattal, 2013; Perone & Hursh, 2013). The five chapters in Part III describe how applied behavior analysts use this experimental approach.
7. Skinner, considered by many the most eminent psychologist of the 20th century (Haagbloom et al., 2002), authored or coauthored 291 primary-source works and a three-volume autobiography (*Particulars of My Life,* 1976; *The Shaping of a Behaviorist,* 1979; *A Matter of Consequences,* 1983). Many of Skinner's books are available as name-your-price products at the B. F. Skinner Foundation (bfskinner.org). Numerous biographical books and articles have been written about Skinner, both before and after his death. *Burrhus Frederic Skinner (1904–1990): A Thank You* by Fred Keller (1976), *B. F. Skinner—The Last Few Days* by his daughter Julie Vargas (1990), *B. F. Skinner, Organism* by Charles Catania (1992), *B. F. Skinner: A Life* by Daniel Bjork (1997), *Skinner as Self-Manager* by Robert Epstein (1997), *Burrhus F. Skinner: Shaper of Behaviour* by Frederik Toates (2009), and a series of articles in the 2017, Quarter 1 issue of *Operants* provide interesting and informative insights about Skinner and his work. Morris, Altus, and Smith (2005) detail Skinner's contributions to ABA. Smith and Morris (2018) maintain a continually updated bibliography of citations, references, and other materials related to Skinner's writings.
8. The nature and role of pragmatism in radical behaviorism are long-running topics of discussion and debate (e.g., Barnes-Holmes, 2000; Baum, 2017; Moore, 2008; Schoneberger, 2016; Tourinho & Neno, 2003; Zuriff, 1980).
9. Excellent discussions and spirited debates about the meaning and usefulness of radical behaviorism as the philosophy for a science of behavior can be found in Baum (2011, 2017); Catania and Harnad (1988); Chiesa (1994); Diller and Lattal (2008); Johnston (2013); Moore (2008, 2011); Palmer (2011); and Rachlin (2018).
10. Boyle and Greer (1983) published an extension of Fuller's classic study with comatose patients.
11. Sidney Bijou's remarkable career and the numerous ways he contributed to the founding and advancement of behavior analysis are detailed with a personal touch by Morris (2008, 2009).

12. Articles describing the histories of the applied behavior analysis programs at five of these universities can be found in the winter 1993 issue of *JABA*. Kenneth Goodall's article, *The Behavior Shapers* (1972), provides a fascinating look at the work of some of the field's pioneers.
13. Furman and Lepper (2018) take issue with this definition of applied behavior analysis because it "employs a subjective criterion (i.e., social relevance and/or importance to man and society), taking it out of the realm of an objective science" (p. 103) and suggest that ABA be defined as "the scientific study of behavior change, using the principles of behavior, to evoke or elicit a targeted behavioral change" (p. 104). We believe *socially significant behavior* is an essential element of the definition of ABA. The science and the philosophy of behaviorism it embodies purposely seek behavior changes that improve people's quality of life (Kimball & Heward, 1993; Skinner, 1971; 1974). Objective (i.e., scientific) methods for assessing and validating the social importance of behaviors targeted for change and the extent to which behavior change improves participants' quality of life are described in Chapters 3 and 10.
14. Choice making in concurrent schedules, delay discounting, and operant variability are discussed in Chapters 13, 27, and 29, respectively.
15. For perspectives on the status, role, and future of basic research in behavior analysis, see Killeen (2018), Marr (2017), and a special series of articles invited by Li, Mahoney, and Poling (2018).
16. To learn more about translational research, see Critchfield (2011a, b) and articles in the Fall 2009 and Spring 2011 issues of *The Behavior Analyst*.

CHAPTER 2

Basic Concepts and Principles

LEARNING OBJECTIVES

- Define behavior, response, and response class.
- State examples of behavior, response, and response class.
- Define stimulus and stimulus class.
- Define and state examples of positive reinforcement.
- Define and state examples of negative reinforcement.
- Define and provide examples of conditioned and unconditioned reinforcement.
- Define and state examples of positive punishment.
- Define and state examples of negative punishment.
- Define and provide examples of stimulus control.
- Define and provide examples of establishing operations.
- Describe a behavioral contingency.
- Describe the respondent conditioning paradigm.
- Provide an example of the respondent conditioning paradigm.
- Describe the operant conditioning paradigm.
- Provide an example of the operant conditioning paradigm.

> One must understand how the organism is modified by its interactions with the environment, how new environment–behavior relations are learned and unlearned.
>
> —Jack L. Michael (2004, p. 1)

This chapter defines the basic concepts required for a scientific analysis of behavior and introduces fundamental principles discovered by such analyses. Behavior, the first concept we examine, is the most fundamental of all. Because the controlling variables of primary importance to applied behavior analysts are located in the environment, the concepts of environment and stimulus are defined next. We then introduce several essential findings that the scientific study of behavior–environment relations has discovered. Two types of behavior—respondent and operant—are described, and the basic ways the environment influences each type—respondent conditioning and operant conditioning—are introduced. The three-term contingency—a concept for expressing temporal and functional relations between operant behavior and environment—and its importance as a focal point in applied behavior analysis are then explained.[1] The chapter's final section recognizes the enormous complexity of human behavior, reminds us that behavior analysts possess an incomplete, though ever-expanding and progressively sophisticated, knowledge of their subject matter, and identifies some of the obstacles and challenges faced by those who strive to change behavior in applied settings.

BEHAVIOR

What, exactly, is behavior? In a nutshell, behavior is the activity of living organisms. Human behavior is everything people do, including how they move and what they say, think, and feel. Tearing open a bag of peanuts is behavior, and so is thinking how good the peanuts will taste once the bag is open. Reading this sentence is behavior, and if you're holding the book, so is feeling its weight and shape in your hands.

Although words such as *activity* and *movement* communicate the general notion of behavior, a more precise definition is needed for scientific purposes. How a scientific discipline defines its subject matter exerts profound influence on the methods of measurement, experimentation, and theoretical analyses that are appropriate and possible.

Building on Skinner's (1938) definition of behavior as "the movement of an organism or of its parts in a frame of reference provided by the organism or by various external objects or fields" (p. 6), Johnston and Pennypacker (1980, 1993, 2009) articulated the most conceptually sound and empirically complete definition of behavior to date. In *Strategies and Tactics of Behavioral Research*, third edition, these authors define **behavior** as follows:

> Behavior is that portion of an organism's interaction with its environment that involves movement of some part of the organism (2009, p. 31).

Johnston and Pennypacker discuss each part of this definition as it relates to researchers and practitioners. The phrase *of an organism* restricts the subject matter to the activity of living organisms, leaving notions such as the "behavior" of the stock market outside the realm of the scientific use of the term. The phrase *an organism's interaction with its environment* "avoids implying that behavior is a possession of the organism and highlights the requirement for an interactive condition" (2009, p. 31). The authors elaborated on this critical part of the definition in the second edition of their text:

> Behavior is not a property or attribute of the organism. It happens only when there is an interactive condition between an organism and its surroundings, which include its own body. This means that independent states of the organism, whether real or hypothetical, are not behavioral events, because there is no interactive process. *Being hungry* or *being anxious* are examples of states that are sometimes confused with the behavior that they are supposed to explain. Neither phrase specifies an environmental agent with which the hungry or anxious organism interacts, so no behavior is implied.
>
> Similarly, independent conditions or changes in the environment do not define behavioral occurrences because no interaction is specified. Someone walking in the rain gets wet, but "getting wet" is not an instance of behavior. A child may receive tokens for correctly working math problems, but "receiving a token" is not behavior. Receiving a token implies changes in the environment but does not suggest or require change in the child's movement. In contrast, both doing math problems and putting the token in a pocket are behavioral events because the environment both prompts the child's actions and is then changed by them. (Johnston & Pennypacker, 1993, p. 24, emphasis added)

In addition to excluding static states of the organism, the definition does not include bodily movements produced by independent physical forces as behavioral events. For example, being blown over by a strong gust of wind is not behavior; given sufficient wind, nonliving objects and organisms move similarly.[2]

The phrase *movement of some part of the organism* identified behavior as movement, regardless of scale. "To be observed, a response must affect the environment—it must have an effect upon an observer or upon an instrument, which in turn can affect an observer. This is as true of the contraction of a small group of muscle fibers as of pressing a lever or pacing a figure 8" (Skinner, 1969, p. 130). Behavior includes movement of body parts within the skin not accessible for observation by others. As Catania (2013) noted, shifting one's attention to different instruments while listening to recorded music need not involve the movement of the head or eyes or any other obvious body part.

The word *behavior* is usually used in reference to a class of responses sharing certain functions (e.g., eating behavior, greeting behavior, writing behavior).[3] The term *response* refers to a specific instance of behavior. A technical definition of **response** is an "*action of an organism's effector.* An effector is an organ at the end of an efferent nerve fiber that is specialized for altering its environment mechanically, chemically, or in terms of other energy changes" (Michael, 2004, p. 8, italics in original). Human effectors include the striated muscles (i.e., skeletal muscles such as biceps and quadriceps), smooth muscles (e.g., stomach and bladder muscles), and glands (e.g., adrenal and pituitary glands).

Behavior can also be described by its form, or physical characteristics. *Response topography* refers to the physical shape or form of behavior. For example, the hand and finger movements used to open a bag of peanuts can be described by their topographical elements. However, careful observation will reveal that the topography differs somewhat each time a person opens a bag of snacks. The difference may be slight, but each "bag opening response" will vary somewhat from all others.

Although describing behavior by its topography is sometimes useful, behavior analysis is characterized by a *functional analysis* of the effects of behavior on the environment. A **response class** is a group of responses with the same function (that is, each response in the group produces the same effect on the environment). Some response classes comprise responses of widely varying form—imagine the tremendous variability in responses to the request to "do something unpredictable" (Neuringer, 2009)—whereas the topographical variation among members of other response classes is limited (e.g., a person's signature, proper grip for a four-seam fastball).

Another reason underscoring the importance of a functional analysis of behavior over a structural or topographical description is that two responses of the same topography can be completely different behaviors depending on the controlling variables. For example, saying "fire" while looking at the letters, *f-i-r-e,* is a vastly different behavior from yelling "Fire!" when smelling smoke or seeing flames in a crowded theater.

Behavior analysts use the term **repertoire** in at least two ways. *Repertoire* is sometimes used to refer to all of the behaviors a person can do. More often the term denotes a person's collection of knowledge and skills relevant to particular settings or tasks. In the latter sense, each of us has learned multiple repertoires. For example, each of us has a repertoire of behaviors appropriate for informal social situations that differs somewhat (or a lot) from the behaviors we use to navigate formal situations. And each person has repertoires with respect to language skills, academic tasks, everyday routines, recreation, and so on. After studying this text, your repertoire of knowledge and skills in applied behavior analysis will be advanced.

ENVIRONMENT

All behavior occurs within an environmental context; behavior cannot be emitted in an environmental void or vacuum. Johnston and Pennypacker (2009) defined environment and noted the implications of that definition for a science of behavior as follows:

> **Environment** refers to the full set of physical circumstances in which the organism exists. The term is comprehensive in that any facets of the physical world may be considered for their contribution to behavior. The term is specific in that for any particular behavior, the focus is usually on only those environmental events that are functionally related to individual responses. . . .

> The relevant environment can even include the organism doing the behaving. . . . [W]hen you scratch an itch, the stimulation from your skin probably increases the effectiveness of relief from the itching as a negative reinforcer and makes the behavior of scratching more likely. Our bodies are an ongoing source of antecedent and consequent environmental events related to responding. This fact reminds us that the skin is not an especially important boundary in understanding of behavior. (p. 29)

The environment is a complex, dynamic universe of events that differs from moment to moment. When behavior analysts describe particular aspects of the environment, they talk in terms of stimulus conditions or events. A **stimulus** is "an energy change that affects an organism through its receptor cells" (Michael, 2004, p. 7). Humans have receptor systems that detect stimulus changes occurring outside and inside the body. *Exteroceptors* are sense organs that detect external stimuli and enable vision, hearing, olfaction, taste, and cutaneous touch. Two types of sense organs sensitive to stimulus changes within the body are *interoceptors,* which are sensitive to stimuli originating in the viscera (e.g., feeling a stomach ache), and *proprioceptors,* which enable the kinesthetic and vestibular senses of movement and balance. Applied behavior analysts most often study the effects of stimulus changes that occur outside the body. External stimulus conditions and events are not only more accessible to observation and manipulation than are internal conditions, but also are key features of the physical and social world in which people live.

The environment influences behavior primarily by stimulus change and not static stimulus conditions. As Michael (2004) noted, when behavior analysts speak of the presentation or occurrence of a stimulus, they usually mean stimulus change.

> For example, in respondent conditioning the conditioned stimulus may be referred to as a tone. However, the relevant event is actually a change from the absence of tone to the tone sounding . . . , and although this is usually understood without having to be mentioned, it can be overlooked in the analysis of more complex phenomena. Operant discriminative stimuli, conditioned reinforcers, conditioned punishers, and conditioned motivative variables are also usually important as stimulus changes, not static conditions (Michael, 2004, pp. 7–8).[4]

Box 2.1, "The Behavioral Stream," illustrates the continuous, ever-changing nature of behavior and the environment.

Classifying and Describing Stimuli

Stimulus events can be described formally (by their physical features), temporally (by when they occur with respect to a behavior of interest), and functionally (by their effects on behavior). Behavior analysts used the term **stimulus class** to refer to any group of stimuli sharing a predetermined set of common elements in one or more of these dimensions.

Formal Dimensions of Stimuli

Behavior analysts often describe, measure, and manipulate stimuli according to their formal dimensions, such as size, color, intensity, weight, and spatial position relative to other objects. Stimuli can be nonsocial (e.g., a red light, a high-pitched sound) or socially mediated (e.g., a friend asking, "Want some more peanuts?").

Temporal Loci of Stimuli

Because behavior and the environmental conditions that influence it occur within and across time, the temporal location of stimulus changes is important. In particular, behavior is affected most by stimulus changes that occur prior to and soon after the behavior. The term **antecedent** refers to environmental conditions or stimulus changes that exist or occur prior to the behavior of interest.

Because behavior cannot occur in an environmental void or vacuum, every response takes place in the context of a particular situation or set of antecedent conditions. These antecedent events play a critical part in learning and motivation, and they do so irrespective of whether the learner or someone in the role of behavior analyst or teacher has planned or is even aware of them.

> For example, just some of the functionally relevant antecedents for a student's performance on a timed math test might include the following: the amount of sleep the student had the night before; the temperature, lighting, and seating arrangements in the classroom; the teacher reminding the class that students who beat their personal best scores on the test will get a free homework pass; and the specific type, format, and sequence of math problems on the test. Each of those antecedent variables (and others) has the potential to exert a great deal, a little, or no noticeable effect on performance as a function of the student's experiences with respect to a particular antecedent. (Heward & Silvestri, 2005, p. 1135)

A **consequence** is a stimulus change that follows a behavior of interest. Some consequences, especially those that are relevant to current motivational states and follow the behavior closely in time, have significant influence on future behavior; other consequences have little effect.

Like antecedent stimulus events, consequences may also be nonsocial events or socially mediated. In a **socially mediated contingency**, another person presents an antecedent stimulus and/or the consequence for the behavior. Table 2.1 on page 47 shows examples of various combinations of nonsocial and socially mediated antecedent and consequent events for four behaviors.

Multiple Functions of Single Stimulus Changes

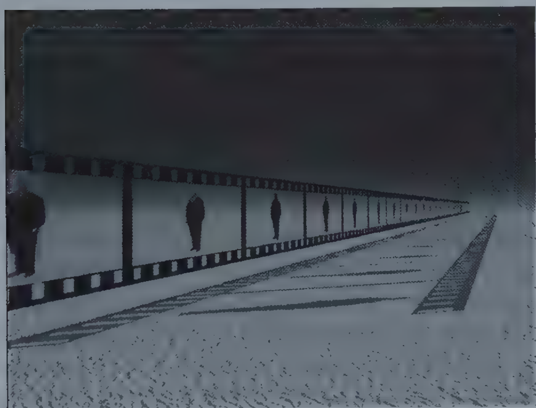
Some stimulus changes exert immediate and powerful control over behavior, whereas others have delayed effects, or no apparent effect. Even though we can and often do describe stimuli by their physical characteristics (e.g., the pitch and decibel level of a tone, the topography of a person's hand and arm movements), stimulus changes are understood best through a functional analysis of their effects on behavior. For example, the same decibel tone that functions in one environment and set of conditions as a prompt for checking the clothes in the dryer may function as a warning signal to fasten a seat belt in another setting or situation; the same hand and arm motion that produces a smile and a "Hi" from another person in one set of conditions occasions a scowl and obscene gesture in another.

Stimulus changes can have one or both of two kinds of basic functions or effects on behavior: (a) an immediate but

BOX 2.1

The Behavioral Stream

Although the concepts of stimulus and response have proven useful for conceptual, experimental, and applied analyses of behavior, it is important to recognize that stimuli and responses do not exist as discrete events in nature. The stimuli and responses identified by scientists and practitioners are detectable "slices" of the continuous, ever-dynamic interaction between an organism and its environment. Mickey Keenan and Karola Dillenburger (2013) call this process the "behavioral stream." They illustrate and describe it as follows:

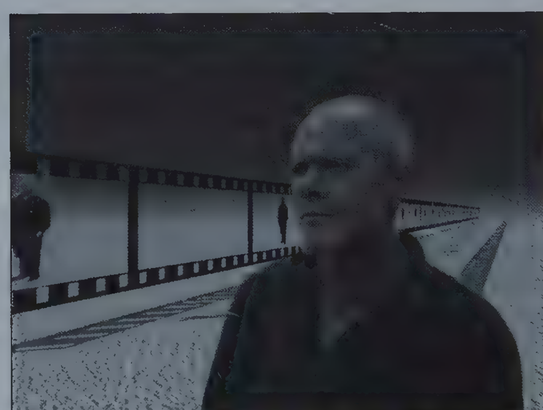


This figure shows the life of an individual to be a continuous process.

Upon first meeting a person, we observe only a snapshot in his life.

The image of a person as a continuous process is difficult to retain.

A consequence of this is that we usually engage in a mentalistic analysis. That is, we attempt to explain a person's behavior by referring to something taking place inside him, in his brain or in his mind.



A natural science perspective, however, enables us to retain the image of a person as a continuous process. From this perspective, the purpose of scientific enquiry is to relate segments of this continuous process to the independent variables that produce them. So-called mental events are part of the continuous process, not the explanation for what is observed.

TABLE 2.1 Antecedent (Situation) and Consequent Events Can Be Nonsocial (Italicized), Socially Mediated (Boldface), or a Combination of Social and Nonsocial

Environmental situation	Response	Consequence
Drink machine	Deposit coins	*Cold drink*
Five cups on table	"One-two-three-four-five cups"	***Teacher nods and smiles***
Friend says "Turn left"	Turn left	*Arrive at destination*
Friend asks "What time is it?"	"Six-fifteen"	***Friend says "Thanks"***

From "Individual Behavior, Culture, and Social Change" by S. S. Glenn, 2004, *The Behavior Analyst, 27*, p. 136. Copyright 2004 by the Association for Behavior Analysis. Used by permission.

temporary effect of increasing or decreasing the frequency of the behavior and/or (b) a delayed but relatively permanent effect in terms of the frequency of that type of behavior in the future (Michael, 1995). For example, a sudden downpour on a cloudy day is likely to increase immediately the occurrence of all behavior that has resulted in escape from rain in the past, such as deploying an umbrella or running for cover under an awning. A person caught in the rain without an umbrella may be more likely to bring one on cloudy days in the future.

RESPONDENT BEHAVIOR

All biologically intact organisms enter the world able to respond in predictable ways to certain stimuli; no learning is required. These ready-made behaviors protect against harmful stimuli (e.g., eyes watering and blinking to remove particles on the cornea), help regulate the internal balance and economy of the organism (e.g., changes in heart rate and respiration in response to changes in temperature and activity levels), and promote reproduction (e.g., sexual arousal). Each of these stimulus–response relations, called a **reflex**, is part of the organism's genetic endowment, a product of natural evolution because of its survival value to the species. At birth, each intact member of a given species comes equipped with the same repertoire of unconditioned (or unlearned) reflexes. Reflexes provide the organism with a set of built-in responses to specific stimuli; these are behaviors the individual organism would not have time to learn. Table 2.2 shows examples of unconditioned reflexes common to humans.

The response component of the stimulus–response reflex is called respondent behavior. **Respondent behavior** is behavior that is elicited by antecedent stimuli. Respondent behavior is induced, or brought out, by a stimulus that precedes the

TABLE 2.2 Examples of Unconditioned Reflexes in Humans

Unconditioned Stimulus	Unconditioned Response	Type of Effector
Loud sound or touch to cornea	Eye blink (lid closes)	Striped muscle
Tactile stimulus under lid or chemical irritant (smoke)	Lacrimal gland secretion (eyes watering)	Gland
Irritation to nasal mucosa	Sneezing	Striped and smooth muscle
Irritation to throat	Coughing	Striped and smooth muscle
Low temperature	Shivering, surface vasoconstriction	Striped and smooth muscle
High temperature	Sweating, surface vasodilation	Gland, smooth muscle
Loud sound	Contraction of tensor tympani and stapedius muscles (reduces amplitude of eardrum vibrations)	Striped muscles
Food in mouth	Salivation	Gland
Indigestible food in stomach	Vomiting	Striped and smooth muscle
Pain stimulus to hand or foot	Hand or foot withdrawal	Striped muscle
A single stimulus that is painful or very intense or very unusual	Activation syndrome—all of the following:	
	Heart rate increase	Cardiac muscle
	Adrenaline secretion	Gland
	Liver release of sugar into bloodstream	Gland
	Constriction of visceral blood vessels	Smooth muscle
	Dilation of blood vessels in skeletal muscles	Smooth muscle
	Galvanic skin response (GSR)	Gland
	Pupillary dilation (and many more)	Smooth muscle

From *Concepts and Principles of Behavior Analysis* (rev. ed.) by J. L. Michael, 2004, pp. 10–11. Copyright 2004 by Society for the Advancement of Behavior Analysis, Kalamazoo, MI.

behavior; nothing else is required for the response to occur. For example, bright light in the eyes (antecedent stimulus) will elicit pupil contraction (respondent). If the relevant body parts (i.e., receptors and effectors) are intact, pupil contraction will occur every time. However, if the eliciting stimulus is presented repeatedly over a short span of time, the strength or magnitude of the response will diminish, and in some cases the response may not occur at all. This process of gradually diminishing response strength is known as **habituation**.

Respondent Conditioning

Formerly neutral stimuli can acquire the ability to elicit respondents through a learning process called **respondent conditioning** (also called *Pavlovian conditioning* and *classical conditioning*). While studying the digestive system of dogs, the Russian physiologist Ivan Petrovich Pavlov (1849–1936) noticed that the animals salivated every time his laboratory assistant opened the cage door to feed them. Dogs do not naturally salivate at the sight of someone in a lab coat, but in Pavlov's laboratory they consistently salivated when the door was opened. His curiosity aroused, Pavlov (1927) designed and conducted an historic series of experiments.[5]

Pavlov started a metronome just an instant before feeding the dogs. Prior to being exposed to this **stimulus–stimulus pairing** procedure, food in the mouth, an **unconditioned stimulus (US)**, elicited salivation, but the sound of the metronome, a **neutral stimulus (NS)**, did not. After several trials consisting of the sound of the metronome followed by the presentation of food, the dogs began salivating in response to the sound of the metronome. The metronome had thus become a **conditioned stimulus (CS)**, and a **conditioned reflex** was established.[6] Respondent conditioning is most effective when the NS is presented just before or simultaneous with the US. However, some conditioning effects can sometimes be achieved with considerable delay between the onset of the NS and the onset of the US, and even with backward conditioning in which the US precedes the NS.

Respondent Extinction

Pavlov also discovered that once a conditioned reflex was established, it would weaken and eventually cease altogether if the conditioned stimulus was presented repeatedly in the absence of the unconditioned stimulus. For example, if the sound of the metronome was presented repeatedly without being accompanied or followed by food, it would gradually lose its ability to elicit salivation. The procedure of repeatedly presenting a conditioned stimulus without the unconditioned stimulus until the conditioned stimulus no longer elicits the conditioned response is called **respondent extinction**.

Figure 2.1 shows schematic representations of respondent conditioning and respondent extinction. In this example, a puff

Before Conditioning	US (air puff)	→	UR (eye blink)
	NS (clicking sound)	----------	no eye blink
Respondent Conditioning	NS + US (click & air puff)	→	UR (eye blink)
	⇓		
	NS + US (click & air puff)	→	UR (eye blink)
	⇓ (more trials)		
Product of Respondent Conditioning	US (air puff)	→	UR (eye blink)
	CS (clicking sound)	→	CR (eye blink)
Respondent Extinction	CS (clicking sound)	→	CR (eye blink)
	⇓		
	CS (clicking sound)	→	CR (eye blink)
	⇓		
	CS (clicking sound)	→	CR (eye blink)
	⇓ (more trials)		
Results of Respondent Extinction	US (air puff)	→	UR (eye blink)
	NS (clicking sound)	----------	no eye blink

Figure 2.1 Schematic representation of respondent conditioning and respondent extinction. The top panel shows an unconditioned reflex: A puff of air (unconditioned stimulus, or US) elicits an eye blink (an unconditioned response, or UR). Before conditioning, a clicking sound (a neutral stimulus, or NS) has no effect on eye blinking. Respondent conditioning consists of a stimulus–stimulus pairing procedure in which the clicking sound is presented repeatedly just prior to, or simultaneously with, the air puff. The product of respondent conditioning is a conditioned reflex (CR): In this case the clicking sound has become a conditioned stimulus (CS) that elicits an eye blink when presented alone. The bottom two panels illustrate the procedure and outcome of respondent extinction: Repeated presentations of the CS alone gradually weaken its ability to elicit eye blinking to the point where the CS eventually becomes an NS again. The unconditioned reflex remains unchanged before, during, and after respondent conditioning.

of air produced by a glaucoma-testing machine is the US for the eye blink reflex. The ophthalmologist's finger pressing the button of the machine makes a faint clicking sound. But prior to conditioning, the clicking sound is an NS: It has no effect on eye blinking. After being paired with the air puff just a few times, the finger-on-the-button sound becomes a CS: It elicits eye blinking as a conditioned reflex.

Conditioned reflexes can also be established by stimulus–stimulus pairing of an NS with a CS. This form of respondent conditioning is called **higher-order** (or *secondary*) **conditioning**. For example, secondary respondent conditioning could occur in a patient who has learned to blink at the clicking sound of the button during the glaucoma-testing situation as follows. The patient detects a slight movement of the ophthalmologist's finger (NS) just before it contacts the button that makes the clicking sound (CS). After several NS–CS pairings, movement of the ophthalmologist's finger may become a CS capable of eliciting blinking.[7]

The form, or topography, of respondent behaviors changes little, if at all, during a person's lifetime. There are two exceptions: (1) Certain reflexes disappear with maturity, such as that of grasping an object placed in the palm of the hand, a reflex usually not seen after the age of 3 months (Bijou & Baer, 1965); and (2) several unconditioned reflexes first appear later in life, such as those related to sexual arousal and reproduction. However, during a person's lifetime an infinite range of stimuli that were previously neutral (e.g., the high-pitched whine of the dentist's drill) can come to elicit respondents (e.g., increased heartbeat and perspiration).

Respondents make up a small percentage of the behaviors typically of interest to applied behavior analysts. As Skinner (1953) pointed out, "Reflexes, conditioned or otherwise, are mainly concerned with the internal physiology of the organism. We are most often interested, however, in behavior which has some effect upon the surrounding world" (p. 59). It is this latter type of behavior, and the process by which it is learned, that we will now examine.

OPERANT BEHAVIOR

A baby in a crib moves her hands and arms, setting in motion a mobile dangling above. The baby is literally operating on her environment, and the mobile's movement and musical sounds—stimulus changes produced by the baby's batting at the toy with her hands—are consequences of her behavior. Her movements are continuously changing as a result of those consequences.

Members of a species whose only way of interacting with the world is a genetically determined fixed set of responses would find it difficult to survive, let alone thrive, in a complex environment that differed from the environment in which their distant ancestors evolved. Although respondent behavior comprises a critically important set of "hard-wired" responses, respondent behavior does not provide an organism with the ability to learn from the consequences of its actions. An organism whose behavior is unchanged by its effects on the environment will be unable to adapt to a changing environment.

Fortunately, in addition to her repertoire of genetically inherited respondent behaviors, our baby entered the world with some *uncommitted behavior* that is highly malleable and susceptible to change through its consequences. This type of behavior, called operant behavior, enables the baby over the course of her life to learn novel, increasingly complex responses to an ever-changing world.[8]

Operant behavior is any behavior determined primarily by its history of consequences. Unlike respondent behavior, which is elicited by antecedent events, operant behavior is selected, shaped, and maintained by the consequences that have followed it in the past.

Unlike respondent behaviors, whose topography and basic functions are predetermined, operant behaviors can take a virtually unlimited range of forms. The form and function of respondent behaviors are constant. By comparison, however, the "meaning" of operant behavior cannot be determined by its topography. Operants are defined functionally, by their effects. Not only does the same operant often include responses of widely different topographies (e.g., a diner may obtain a glass of water by nodding his head, pointing to a glass of water, or saying yes to a waiter), but also, as Skinner (1969) explained, the same movements comprise different operants under different conditions.

> Allowing water to pass over one's hands can perhaps be adequately described as topography, but "washing one's hands" is an "operant" defined by the fact that, when one has behaved this way in the past, one's hands have become clean—a condition which has become reinforcing because, say, it has minimized a threat of criticism or contagion. Behavior of precisely the same topography would be part of another operant if the reinforcement had consisted of simple stimulation (e.g., "tickling") of the hands or the evocation of imitative behavior in a child whom one is teaching to wash his hands. (p. 127)

Topography is of little or no use in identifying behavior as respondent or operant. For example, crying elicited by painful stimuli is respondent behavior, but the same tear production is operant when a function of parental attention (cf., Bowman, Hardesty, & Mendres-Smith, 2013; Epstein, 2012).

Table 2.3 compares and contrasts defining features and key characteristics of respondent behavior and operant behavior.

Selection by Consequences

> *Human behavior is the joint product of (i) the contingencies of survival responsible for the natural selection of the species and (ii) the contingencies of reinforcement responsible for the repertoires acquired by its members, including (iii) the special contingencies maintained by the social environment. [Ultimately, of course, it is all a matter of natural selection, since operant conditioning is an evolved process, of which cultural practices are special applications.]*
>
> —B. F. Skinner (1981 p. 502)

Skinner's discovery and subsequent elucidation of operant selection by consequences have rightly been called "revolutionary" and

TABLE 2.3 Comparing and Contrasting Defining Features and Key Characteristics of Respondent and Operant Behavior

Characteristics or features	Respondent behavior	Operant behavior
Definition	Behavior elicited by antecedent stimuli.	Behavior selected by its consequences.
Basic unit	Reflex: an antecedent stimulus elicits a particular response (S–R).	Operant response class: A group of responses each of which produces the same effect on the environment; described by three-term contingency relation of antecedent stimulus conditions, behavior, and consequence (A–B–C).
Examples	Newborn's grasping and suckling to touch; pupil constriction to bright light; cough/gag to irritation in throat; salivation at smell of food; withdrawing hand from painful stimulus; sexual arousal to stimulation.	Talking, walking, playing the piano, riding a bike, counting change, baking a pie, hitting a curveball, laughing at a joke, thinking about a grandparent, reading this book.
Body parts (effectors) that most often produce the response (not a defining feature)	Primarily smooth muscles and glands (adrenaline squirt); sometimes striated (skeletal) muscles (e.g., knee-jerk to tap just below patella).	Primarily striated (skeletal) muscles; sometimes smooth muscles and glands.
Function or usefulness for individual organism	Maintains internal economy of the organism; provides a set of "ready-made" survival responses the organism would not have time to learn.	Enables effective interaction and adaptation in an ever-changing environment that could not be anticipated by evolution.
Function or usefulness for species	Promotes continuation of species indirectly (protective reflexes help individuals survive to reproductive age) and directly (reflexes related to reproduction).	Individuals whose behavior is most sensitive to consequences are more likely to survive and reproduce.
Conditioning process	Respondent (also called classical or Pavlovian) conditioning: Through a stimulus–stimulus pairing procedure in which a neutral stimulus (NS) presented just prior to or simultaneous with an unconditioned (US) or conditioned (CS) eliciting stimulus, the NS becomes a CS that elicits the response and a conditioned reflex is created. (See Figure 2.1)	Operant conditioning: Some stimulus changes following a response closely in time increase (reinforcement) or decrease (punishment) future occurrences of similar responses under similar conditions. Previously neutral stimulus changes become conditioned reinforcers or punishers as result of stimulus–stimulus pairing with other reinforcers or punishers.
Repertoire limits	Topography and function of respondents determined by natural evolution of species (phylogeny). All biologically intact members of a species possess the same set of unconditioned reflexes. Although new forms of respondent behavior are not learned, an infinite number of conditioned reflexes may emerge in an individual's repertoire, depending on the stimulus–stimulus pairing the individual has experienced (ontogeny).	Topography and function of each person's repertoire of operant behaviors are selected by consequences during the individual's lifetime (ontogeny). New and more complex operant response classes can emerge. Response products of some human operants (e.g., airplanes) enable some behaviors not possible by anatomical structure alone (e.g., flying).

"the bedrock on which other behavioral principles rest" (Glenn, 2004, p. 134). **Selectionism** "anchors a new paradigm in the life sciences. . . . A basic tenet of this position is that all forms of life, from single cells to complex cultures, evolve as a result of selection with respect to function" (Pennypacker, 1994, pp. 12–13).

Selection by consequences operates during the lifetime of the individual organism (**ontogeny**) and is a conceptual parallel to Darwin's (1872/1958) natural selection in the evolutionary history of a species (**phylogeny**).[9] In response to the question "Why do giraffes have long necks?" Baum (2017) provides this excellent description of natural selection:

> Darwin's great contribution was to see that a relatively simple mechanism could help explain why phylogeny followed the particular course it did. The history of giraffes' necks, Darwin saw, is more than a sequence of changes; it is a history of selection. What does the selecting? Not an omnipotent Creator, not Mother Nature, not the giraffes, but a natural, mechanical process: natural selection.
>
> Within any population of organisms, individuals vary. They vary partly because of environmental factors (e.g., nutrition), and also because of genetic inheritance. Among the giraffe ancestors that lived in what is now the Serengeti Plain, for instance, variation in genes meant that some had shorter necks and some had longer necks.
>
> As the climate gradually changed, however, new, taller types of vegetation became more frequent. The giraffe ancestors that had longer necks, being able to reach higher,

got a little more to eat, on the average. As a result, they were a little healthier, resisted disease a little better, evaded predators a little better—on the average. Any one individual with a longer neck may have died without offspring, but on the average longer-necked individuals had more offspring, which tended on the average to survive a little better and produce more offspring. As longer necks became frequent, new genetic combinations occurred, with the result that some offspring had still longer necks than those before, and they did still better. As the longer-necked giraffes continued to out-reproduce the shorter-necked ones, the population consisted more and more of longer-necked individuals, and the average neck length of the whole population grew. (pp. 59–60)

Just as natural selection requires a population of individual organisms with varied physical features (e.g., giraffes with necks of different lengths), operant selection by consequences requires variation in behavior. Those behaviors that produce the most favorable outcomes are selected and "survive," which leads to a more adaptive repertoire. Natural selection has endowed humans with an initial population of uncommitted behavior (e.g., babies babbling and moving their limbs about) that is highly malleable and susceptible to the influence of the consequences that follow it. As Glenn (2004) noted,

> By outfitting humans with a largely uncommitted behavioral repertoire, natural selection gave our species a long leash for local behavioral adaptations. But the uncommitted repertoire of humans would be lethal without the . . . susceptibility of human behavior to operant selection. Although this behavioral characteristic is shared by many species, humans appear to be most exquisitely sensitive to behavioral contingencies of selection (Schwartz, 1974). (p. 139)

Operant Conditioning

> *Operant conditioning may be seen everywhere in the multifarious activities of human beings from birth until death. . . . It is present in our most delicate discriminations and our subtlest skills; in our earliest crude habits and the highest refinements of creative thought.*
>
> —Keller and Schoenfeld (1950, p. 64)

Operant conditioning refers to the process and selective effects of consequences on behavior.[10] From an operant conditioning perspective, a functional consequence is a stimulus change that follows a given response closely in time and alters the occurrence of similar responses in the future. "In operant conditioning we 'strengthen' an operant in the sense of making a response more probable or, in actual fact, more frequent" (Skinner, 1953, p. 65). If the movement and sounds produced by the baby's batting at the mobile with her hands increase the occurrence of hand movements in the direction of the toy, operant conditioning has taken place.

When operant conditioning consists of an increase in response rate, *reinforcement* has taken place, and the consequence responsible, in this case the movement and sound of the mobile, would be called a **reinforcer**.[11] Although operant conditioning is used most often to refer to the "strengthening" effects of reinforcement, as Skinner described earlier, it also encompasses the principle of punishment. If the mobile's movement and musical sounds resulted in a decrease in the rate at which the baby bats it with her hands, *punishment* has occurred, and the mobile's movement and sound would be called **punishers**. Before we examine the principles of reinforcement and punishment further, it is important to identify several important qualifications concerning how consequences affect behavior.

Consequences Affect Only Future Behavior

Consequences affect only future behavior. Specifically, a behavioral consequence affects the relative rate at which similar responses will be emitted in the future under similar stimulus conditions. This point may seem too obvious to merit mention because it is both logically and physically impossible for a consequent event to affect a behavior that preceded it, when that behavior is over before the consequent event occurs. Nevertheless, the statement "behavior is controlled by its consequences" raises the question. (See Box 2.2, "When the Phone Rings," for further discussion of this apparent logical fallacy.)

Consequences Select Response Classes, Not Individual Responses

Responses emitted because of the effects of reinforcement will differ slightly from previously reinforced responses, but will share at least one functional element with the former responses to produce the same consequence.

> Reinforcement strengthens responses which differ in topography from the response reinforced. When we reinforce pressing a lever, for example, or saying Hello, responses differing quite widely in topography grow more probable. This is a characteristic of behavior which has strong survival value . . . , since it would be very hard for an organism to acquire an effective repertoire if reinforcement strengthened only identical responses. (Skinner, 1969, p. 131)

It is the response class—topographically different, but functionally similar responses—that is strengthened or weakened by operant conditioning. "An operant activity is a population of actions all of which have the same environmental effect" (Baum, 2017, p. 84). The concept of response class is "implied when it is said that reinforcement increases the future frequency of the *type* of behavior that immediately preceded the reinforcement" (Michael, 2004, p. 9). As will be shown in later chapters, the concept of response class is a key to the development and elaboration of new behavior.

If consequences (or natural evolution) selected only a very narrow range of responses (or genotypes), the effect would "tend toward uniformity and a perfection of sorts" (Moxley, 2004, p. 110) that would place the behavior (or species) at risk of extinction should the environment change. For example, if the mobile's movement and sound reinforced only arm and hand movements that fell within an exact and narrow range of motion and no similar movements survived, the baby would be unable to contact that reinforcement if one day her mother mounted the mobile in a different location above the crib.

BOX 2.2

When the Phone Rings

The professor was ready to move on to his next point, but a raised hand in the front row caught his attention.

Professor: Yes?

Student: You said that operant behavior, like talking, writing, running, reading, driving a car, most everything we do—you said all of those behaviors are controlled by their consequences, by things that happen *after* the response was emitted?

Professor: I said that. Yes.

Student: Well, I have a hard time with that. When my telephone rings and I pick up the receiver, that's an operant response, right? I mean, answering the phone when it rings certainly didn't evolve genetically as a reflex to help our species survive. So, we're talking about operant behavior, correct?

Professor: Correct.

Student: All right then. How can we say my picking up my telephone is controlled by its consequence? I answer the phone *because* it is ringing. So does everyone. Ringing controls the response. And ringing can't be a consequence because it comes before the response.

The professor hesitated replying just long enough for the student to believe himself the hero, nailing a professor for pontificating about some theoretical concept that has little or no relevance to the everyday real world. Simultaneously sensing victory, other students began to pile on.

Another Student: How about stepping on the brake when you see a stop sign? The sign controls the braking response, and that's not a consequence either.

A Student from the Back of the Room: And take a common classroom example. When a kid sees the problem 2 + 2 on his worksheet and writes 4, writing the numeral 4 has to be controlled by the written problem itself. Otherwise, how could anyone learn the correct answers to any question or problem?

Most of the Class: Yah, that's right!

Professor: (with a wry smile) All of you are correct. . . . So, too. am I.

Someone Else in the Class: What do you mean?

Professor: That was exactly my next point, and I was hoping you would pick up on it. (The professor smiled a thank you at the student who had started the discussion and went on.) All around us, every day, we are exposed to countless changing stimulus conditions. The situations you've described are excellent examples of what behavior analysts call *stimulus control*. When the rate of a given behavior is higher in the presence of a given stimulus than when that stimulus is absent, we say that stimulus control is at work. Stimulus control is a fundamentally important principle in behavior analysis, and it will be the subject of much discussion this semester.

But, and here's the important point: A discriminative stimulus, the antecedent event that comes before the response of interest, acquires its ability to evoke a particular response class because of its pairing with certain consequences in the past. It's not just the phone's ring that causes you to pick up the receiver. It is the fact that in the past, answering the phone when it was ringing was followed by the sound of a person's voice. It's that person talking to you, the consequence of picking up the receiver, that selected phone-answering behavior in the first place, but you only answer ringing phones. Why? Because you have learned that there's someone on the other end only when the phone's ringing. So we can still speak of consequences as having the ultimate control in terms of selecting operant behavior, but antecedent stimuli, by virtue of being paired with differential consequences, can indicate what kind of consequence is likely. This concept is called the three-term contingency, and its understanding, analysis, and manipulation are central to applied behavior analysis.

Immediate Consequences Have the Greatest Effect

Behavior is most sensitive to stimulus changes that occur immediately after, or within a few seconds of, a response.

> It is essential to emphasize the importance of the immediacy of reinforcement. Events that are delayed more than a few seconds after the response do not *directly* increase its future frequency. When human behavior is apparently affected by long-delayed consequences, the change is accomplished by virtue of the human's complex social and verbal history, and should not be thought of as an instance of the simple strengthening of behavior by reinforcement. (p. 110) . . . [As with reinforcement,] the longer the time delay between the occurrence of the response and the occurrence of the stimulus change (between R and S^P), the less effective the punishment will be in changing the relevant response frequency, but not much is known about upper limits. (Michael, 2004, p. 36, emphasis in original, words in brackets added)

Consequences Select Any Behavior

Reinforcement and punishment are "equal opportunity" selectors. No logical or healthy or (in the long run) adaptive connection between a behavior and the consequence that functions to strengthen or weaken it is necessary. Any behavior that

immediately precedes reinforcement (or punishment) will be increased (or decreased).

It is the *temporal relation* between behavior and consequence that is functional, not the topographical or logical ones. "So far as the organism is concerned, the only important property of the contingency is temporal. The reinforcer simply *follows* the response. How this is brought about does not matter" (Skinner, 1953, p. 85, emphasis in original). The arbitrary nature of which behaviors are reinforced (or punished) in operant conditioning is exemplified by the appearance of idiosyncratic behaviors that have no apparent purpose or function. An example is the superstitious routine of a poker player who taps and arranges his cards in a peculiar fashion because similar movements in the past were followed by winning hands.

Operant Conditioning Occurs Automatically

Operant conditioning does not require a person's awareness. "A reinforcing connection need not be obvious to the individual [whose behavior is] reinforced" (Skinner, 1953, p. 75, words in brackets added). This statement refers to the **automaticity of reinforcement**; that is, behavior is modified by its consequences regardless of whether the individual is aware that her behavior is, or has been, reinforced.[12] A person does not have to understand or verbalize the relation between her behavior and a consequence, or even know that a consequence has occurred, for reinforcement to "work."

Reinforcement

Reinforcement is the most important principle of behavior and a key element of most behavior change programs designed by behavior analysts (Flora, 2004; Northup, Vollmer, & Serrett, 1993). When a response is followed by a stimulus change that results in similar responses occurring more often, **reinforcement** has taken place.[13] Sometimes the delivery of just one reinforcer results in significant behavior change, although most often several responses must be followed by reinforcement before significant conditioning will occur.

Most stimulus changes that function as reinforcers can be described operationally as either (a) a new stimulus added to the environment (or increased in intensity), or (b) an already present stimulus removed from the environment (or reduced in intensity).[14] These two operations provide for two forms of reinforcement, called positive and negative (see Figure 2.2).

In **positive reinforcement**, a response is followed immediately by the presentation of a stimulus that results in similar responses occurring more often. Our baby's increased batting of the mobile with her hands, when doing so produces movement and music, is an example of positive reinforcement. Likewise, a child's independent play is reinforced when it increases as a result of his parent's giving praise and attention when he plays. Positive reinforcement and procedures for using it to promote desired behaviors are described in detail in Chapter 11.

When a behavior occurs more often because past responses have resulted in the withdrawal or termination of a stimulus, the operation is called **negative reinforcement**. Skinner (1953) used the term **aversive stimulus** to refer to, among other things, stimulus conditions whose termination functioned as reinforcement. Let us assume now that a parent programs the mobile to automatically play music for a period of time. Let us also assume that if the baby bats the mobile with hands or feet, the music immediately stops for a few seconds. If the baby bats the mobile more often when doing so terminates the music, negative reinforcement is at work, and the music can be called *aversive*.

Negative reinforcement is characterized by escape or avoidance contingencies. Jumping out of the shower when the water suddenly becomes too hot is negatively reinforced by *escape* from the burning water. Being sent to the principal's office for acting out may function as negative reinforcement if it enables the misbehaving student to *avoid* an aversive (to him) classroom activity.

The concept of negative reinforcement has confused many students of behavior analysis. Much of the confusion can be traced to the inconsistent early history and development of the term and to psychology and education textbooks and professors who have used the term inaccurately.[15] The most common mistake is equating negative reinforcement with punishment. To help avoid the error, Michael (2004) suggested the following:

> Think about how you would respond if someone asked you (1) whether or not you like negative reinforcement; also if you were asked (2) which you prefer, positive or negative

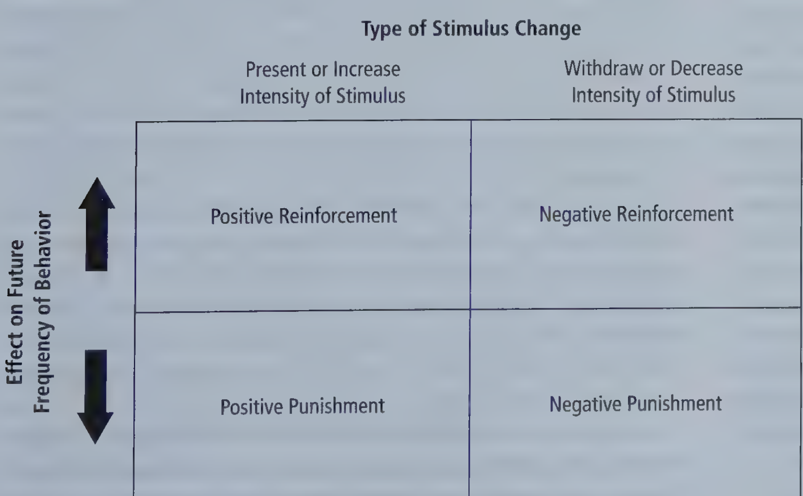


Figure 2.2 Positive and negative reinforcement and positive and negative punishment are defined by the type of stimulus change operation that follows a behavior and the effect that operation has on future occurrences of that type of behavior.

> reinforcement. Your answer to the first question should be that you do indeed like negative reinforcement, which consists of the removal or termination of an aversive condition that is already present. The term *negative reinforcement* refers **only** to the termination of the stimulus. In a laboratory procedure the stimulus must, of course, be turned on and then its termination can be made contingent upon the critical response. No one wants an aversive stimulus turned on, but once it is on, its termination is usually desirable. Your answer to the second question should be that you cannot choose without knowing the specifics of the positive and negative reinforcement involved. The common error is to choose positive reinforcement, but removal of a very severe pain would certainly be preferred over the presentation of a small monetary reward or an edible, unless the food deprivation was very severe. (p. 32, italics and bold type in original)

Negative reinforcement is examined in detail in Chapter 12. Remembering that the term *reinforcement* always means an increase in the occurrence of the behavior and that the modifiers *positive* and *negative* describe the type of stimulus change operation that best characterizes the consequence (i.e., adding or withdrawing a stimulus) should facilitate the discrimination of the principles and application of positive and negative reinforcement.

After a behavior has been established with reinforcement, it need not be reinforced each time it occurs. Many behaviors are maintained at high levels by schedules of *intermittent reinforcement*. Chapter 13 describes various *schedules of reinforcement* and their effects on behavior. However, if reinforcement is withheld for all members of a previously reinforced response class, a procedure based on the principle of **extinction**, the behavior will gradually decrease in rate to its pre-reinforcement level or cease to occur altogether. Chapter 24 describes the principle of extinction and the use of behavior change tactics based on extinction to decrease undesired behavior.

Punishment

Punishment, like reinforcement, is defined functionally. When a response is followed immediately by a stimulus change that results in similar responses occurring less often, **punishment** has taken place. Punishment can be accomplished by either of two types of stimulus change operations (see the bottom two boxes of Figure 2.2). As with reinforcement, the modifiers in the terms *positive punishment* and *negative punishment* connote neither the intention nor the desirability of the behavior change produced; they indicate only that the stimulus change that served as the punishing consequence was presented (**positive punishment**) or withdrawn (**negative punishment**).[16]

As with positive and negative reinforcement, numerous behavior change procedures incorporate the two basic punishment operations. Although some textbooks reserve the term *punishment* for procedures involving positive punishment and describe *time-out from positive reinforcement* and *response cost* as separate "principles" or types of punishment, both methods for reducing behavior are based on negative punishment (see Chapter 15).

Reinforcement and punishment can each be accomplished by either of two different operations, depending on whether the consequence consists of presenting a new stimulus (or increasing the intensity of a current stimulus) or withdrawing (or decreasing the intensity of) a currently present stimulus in the environment (Morse & Kelleher, 1977; Skinner, 1953). Some behavior analysts argue that from a functional and theoretical standpoint only two principles are required to describe the basic effects of behavioral consequences—reinforcement and punishment.[17] However, from a procedural perspective (a critical factor for the applied behavior analyst), a number of behavior change tactics are derived from each of the four operations represented in Figure 2.2.

Most behavior change procedures involve several principles of behavior (see Box 2.3). It is critical for the behavior analyst to have a solid conceptual understanding of the basic principles of behavior. Such knowledge permits better analysis of current controlling variables as well as more effective design and assessment of behavioral interventions that recognize the role various principles may be playing in a given situation.

Stimulus Changes That Function as Reinforcers and Punishers

Because operant conditioning involves the consequences of behavior, it follows that anyone interested in using operant conditioning to change behavior must identify and control the occurrence of relevant consequences. For the applied behavior analyst, therefore, an important question becomes: What kinds of stimulus changes function as reinforcers and punishers?

Unconditioned Reinforcement and Punishment

Some stimulus changes function as reinforcement even though the organism has had no particular learning history with those stimuli. A stimulus change that can increase future occurrences of behavior without prior pairing with any other form of reinforcement is called an **unconditioned reinforcer**.[18] For example, stimuli such as food, water, and sexual stimulation that support the biological maintenance of the organism and survival of the species often function as unconditioned reinforcers. The words *can* and *often* in the two previous sentences recognize the important qualification that the momentary effectiveness of an unconditioned reinforcer is a function of current **motivating operations**. For example, a certain level of food **deprivation** is necessary for the presentation of food to function as a reinforcer. However, food is unlikely to function as reinforcement for a person who has recently eaten a heavy meal. The nature and functions of motivating operations are described in detail in Chapter 16.

Similarly, an **unconditioned punisher** is a stimulus change that can decrease the future occurrence of any behavior that precedes it without prior pairing with any other form of punishment. Unconditioned punishers include painful stimulation that can cause tissue damage (i.e., harm body cells). However, virtually any stimulus to which an organism's receptors are sensitive—light, sound, and temperature, to name a few—can be intensified to the point that its delivery will suppress behavior even though the stimulus is below levels that actually cause tissue damage (Bijou & Baer, 1965).

BOX 2.3

Distinguishing Between Principles of Behavior and Behavior Change Tactics

A principle of behavior describes a basic behavior–environment relation that has been demonstrated repeatedly in hundreds, even thousands, of experiments. A **principle of behavior** describes a functional relation between behavior and one or more of its controlling variables (in the form of $y = fx$) that has thorough generality across individual organisms, species, settings, and behaviors. A principle of behavior is an empirical generalization inferred from many experiments. Principles describe how behavior works. Some examples of principles are reinforcement, punishment, and extinction.

In general, a behavior change tactic is a method for operationalizing, or putting into practice, one or more principles of behavior. A **behavior change tactic** is a research-based, technologically consistent method for changing behavior that has been derived from one or more basic principles of behavior and that possesses sufficient generality across subjects, settings, and/or behaviors to warrant its codification and dissemination. Behavior change tactics constitute the technological aspect of applied behavior analysis. Examples of behavior change tactics include backward chaining, differential reinforcement of other behavior, shaping, response cost, and time-out from positive reinforcement.

So, principles describe basic scientific laws that delineate how behavior–environment relations work, and behavior change tactics are how applied behavior analysts put the principles to work to help people learn and use socially significant behaviors. There are relatively few behavior principles, but many derivative behavior change tactics. To illustrate further, reinforcement is a behavioral principle because it describes a lawful relation between behavior, an immediate consequence, and an increased occurrence of the behavior in the future under similar conditions. However, issuing of checkmarks in a token economy and providing contingent praise are behavior change tactics derived from the principle of reinforcement. To cite another example, punishment is a principle of behavior because it describes the established functional relation between the presentation of a consequence and the decreased occurrence of similar behavior in the future. Response cost and time-out, in contrast, are tactics for changing behavior based on the principle of punishment.

Events that function as unconditioned reinforcers and punishers are the product of the natural evolution of the species (phylogeny). Malott, Tillema, and Glenn (1978) described the natural selection of "rewards" and "aversives" as follows:[19]

> Some rewards and aversives control our actions because of the way our species evolved; we call these unlearned rewards or aversives. We inherit a biological structure that causes some stimuli to be rewarding or aversive. This structure evolved because rewards helped our ancestors survive, while aversives hurt their survival. Some of these unlearned rewards, such as food and fluid, help us survive by strengthening our body cells. Others help our species survive by causing us to produce and care for our offspring—these stimuli include the rewarding stimulation resulting from copulation and nursing. And many unlearned aversives harm our survival by damaging our body cells; such aversives include burns, cuts and bruises. (p. 9)

While unconditioned reinforcers and punishers are critically important and necessary for survival, relatively few behaviors that constitute the everyday routines of people as they go about working, playing, and socializing are directly controlled by such events. For example, although going to work each day earns the money that buys food, eating that food is far too delayed for it to exert any direct operant control over the behavior that earned it. Remember: Behavior is most affected by its immediate consequences.

Conditioned Reinforcers and Punishers

Stimulus events or conditions that are present or that occur just before, or simultaneous with, the occurrence of other reinforcers (or punishers) may acquire the ability to reinforce (or punish) behavior when they later occur on their own as consequences. Called **conditioned reinforcers** and **conditioned punishers**, these stimulus changes function as reinforcers and punishers because of their prior pairing with other reinforcers or punishers.[20] The stimulus–stimulus pairing procedure responsible for the creation of conditioned reinforcers or punishers is the same as that used for respondent conditioning except that the "outcome is a stimulus that functions as a reinforcer [or punisher] rather than a stimulus that will elicit a response" (Michael, 2004, p. 66, words in brackets added).

Conditioned reinforcers and punishers are not related to any biological need or anatomical structure; their ability to modify behavior is a result of each person's unique history of interactions with his or her environment (ontogeny). On the one hand, because no two people experience the world in exactly the same way, the roster of events that can serve as conditioned reinforcers and punishers at any particular time (given a relevant motivating operation) is idiosyncratic to each individual and always changing. On the other hand, to the extent that two people have had similar experiences (e.g., schooling, profession, the culture in general), they are likely to be affected in similar ways by similar events. Social praise and attention are examples of widely effective conditioned reinforcers in many cultures. Because social attention and approval (as well as disapproval) are often paired with many other reinforcers (and punishers), they exert powerful control over human behavior and will be featured in later chapters when specific behavior change tactics are presented.

Because people who live in a common culture share similar histories, it is not unreasonable for a practitioner to search

for potential reinforcers and punishers for a given client among classes of stimuli that have proven effective with other similar clients. However, in an effort to help the reader establish a fundamental understanding of the nature of operant conditioning, we have purposely avoided presenting a list of stimuli that may function as reinforcers and punishers. Morse and Kelleher (1977) made this important point very well.

> Reinforcers and punishers, as environmental "things," appear to have a greater reality than orderly temporal changes in ongoing behavior. Such a view is deceptive. There is no concept that predicts reliably when events will be reinforcers or punishers; *the defining characteristics of reinforcers and punishers are how they change behavior* [italics added]. Events that increase or decrease the subsequent occurrence of one response may not modify other responses in the same way.
>
> In characterizing reinforcement as the presentation of a reinforcer contingent upon a response, the tendency is to emphasize the event and to ignore the importance of both the contingent relations and the antecedent and subsequent behavior. It is *how* [italics added] they change behavior that defines the terms *reinforcer* and *punisher;* thus it is the orderly change in behavior that is the key to these definitions. It is *not* [italics added] appropriate to presume that particular environmental events such as the presentation of food or electric shock are reinforcers or punishers until a change in the rate of responding has occurred when the event is scheduled in relation to specified responses.
>
> A stimulus paired with a reinforcer is said to have become a conditioned reinforcer, but actually it is the behaving subject that has changed, not the stimulus. . . . It is, of course, useful shorthand to speak of conditioned reinforcers . . . just as it is convenient to speak about a reinforcer rather than speaking about an event that has followed an instance of a specific response and resulted in a subsequent increase in the occurrence of similar responses. The latter may be cumbersome, but it has the advantage of empirical referents. Because many different responses can be shaped by consequent events, and because a given consequent event is often effective in modifying the behavior of different individuals, it becomes common practice to refer to reinforcers without specifying the behavior that is being modified. These common practices have unfortunate consequences. They lead to erroneous views that responses are arbitrary and that the reinforcing or punishing effect of an event is a specific property of the event itself. (pp. 176–177, 180)

The point made by Morse and Kelleher (1977) is of paramount importance to understanding functional behavior–environment relations. Reinforcement and punishment are not simply the products of certain stimulus events, which are then called reinforcers and punishers without reference to a given behavior and environmental conditions. There are no inherent or standard physical properties of stimuli that determine their permanent status as reinforcers and punishers. In fact, a stimulus can function as a positive reinforcer under one set of conditions and a negative reinforcer under different conditions. Just as positive reinforcers are not defined with terms such as *pleasant* or *satisfying,* aversive stimuli are not defined with terms such as *annoying* or *unpleasant.* The terms *reinforcer* and *punisher* should not to be used on the basis of a stimulus event's assumed effect on behavior or on the basis of any inherent property of the stimulus event itself. Morse and Kelleher (1977) continued:

> When the borders of the table are designated in terms of stimulus classes (positive–negative; pleasant–noxious) and experimental operations (stimulus presentation–stimulus withdrawal), the cells of the table are, by definition, varieties of reinforcement and punishment. One problem is that the processes indicated in the cells have already been assumed in categorizing stimuli as positive or negative; a second is that there is a tacit assumption that the presentation or withdrawal of a particular stimulus will have an invariant effect. These relations are clearer if empirical operations are used to designate the border conditions. . . . The characterization of behavioral processes depends upon empirical observations. The same stimulus event, under different conditions, may increase behavior or decrease behavior. In the former case the process is called *reinforcement* and in the latter the process is called *punishment.* (p. 180)

At the risk of redundancy, we restate this important concept. Reinforcers and punishers denote functional classes of stimulus events, membership to which is not based on the physical nature of the stimulus changes or events themselves. Indeed, given a person's individual history and current motivational state, and the current environmental conditions, "any stimulus change can be a 'reinforcer' if the characteristics of the change, and the temporal relation of the change to the response under observation, are properly selected" (Schoenfeld, 1995, p. 184). Thus, the phrase "everything is relative" is thoroughly relevant to understanding functional behavior–environment relations.

The Discriminated Operant and Three-Term Contingency

We have discussed the role of consequences in influencing the future occurrence of behavior. But operant conditioning does much more than establish a functional relation between behavior and its consequences. Operant conditioning also establishes functional relations between behavior and certain antecedent conditions.

> In contrast to *if-A-then-B* formulations (such as S-R formulations), the *AB-because-of-C* formulation is a general statement that the relation between an event (B) and its context (A) is because of consequences (C). . . . Applied to Skinner's three-term contingency, the relation between (A) the setting and (B) behavior exists because of (C) consequences that occurred for previous AB (setting-behavior) relations. The idea [is] that reinforcement strengthens the setting-behavior relation rather than simply strengthening behavior. (Moxley, 2004, p. 111)

Reinforcement selects not just certain forms of behavior; it also selects the environmental conditions that in the future will evoke (make more likely) instances of the

response class. A behavior that occurs more often under some antecedent conditions than it does in others is called a **discriminated operant**. Because a discriminated operant occurs more often in the presence of a given stimulus than it does in the absence of that stimulus, it is said to be under **stimulus control**. Answering the phone, one of the everyday behaviors discussed by the professor and his students in Box 2.2, is a discriminated operant. The phone's ring functions as a **discriminative stimulus** ($\mathbf{S^D}$) for answering the phone. We answer the phone when it is ringing, and we do not answer the phone when it is silent.

Just as reinforcers or punishers cannot be identified by their physical characteristics, stimuli possess no inherent dimensions or properties that enable them to function as discriminative stimuli. Operant conditioning brings behavior under the control of various properties or values of antecedent stimuli (e.g., size, shape, color, spatial relation to another stimulus), and what those features are cannot be determined a priori. (Stimulus control is described in detail in Chapter 17.)

> Any stimulus present when an operant is reinforced acquires control in the sense that the rate will be higher when it is present. Such a stimulus does not act as a goad; it does not elicit the response in the sense of forcing it to occur. It is simply an essential aspect of the occasion upon which a response is made and reinforced. The difference is made clear by calling it a discriminative stimulus (or S^D). An adequate formulation of the interaction between an organism and its environment must always specify three things: (1) the occasion upon which a response occurs; (2) the response itself; and (3) the reinforcing consequences. The interrelationships among them are the "contingencies of reinforcement." (Skinner, 1969, p. 7)

The discriminated operant has its origin in the three-term contingency. The **three-term contingency**—*a*ntecedent, *b*ehavior, and *c*onsequence—is sometimes called the ABCs of behavior analysis. Figure 2.3 shows examples of three-term contingencies for positive reinforcement, negative reinforcement, positive punishment, and negative punishment.[21] Most of what the science of behavior analysis has discovered about the prediction and control of human behavior involves the three-term contingency, which is "considered the basic unit of analysis in the analysis of operant behavior" (Glenn, Ellis, & Greenspoon, 1992, p. 1332).

A four-term contingency analysis takes into account motivating events that make certain stimulus changes momentarily more or less reinforcing. The four-term contingency is introduced in Chapter 11 and described in detail in Chapter 16.

The term **contingency** appears in behavior analysis literature with several meanings signifying various types of temporal and functional relations between behavior and antecedent and consequent variables (Catania, 2013; Lattal, 1995; Lattal & Shahan, 1997; Vollmer & Hackenberg, 2001). Perhaps the most common connotation of contingency refers to the *dependency* of a particular consequence on the occurrence of the behavior. When a reinforcer (or punisher) is said to be **contingent** on a particular behavior, the behavior must be emitted for the consequence to occur. For example, after saying, "Name a carnivorous dinosaur," a teacher's "Well done!" depends on the student's response, "Tyrannosaurus Rex" (or another dinosaur of the same class).[22]

The term *contingency* is also used in reference to the *temporal contiguity* of behavior and its consequences. As stated

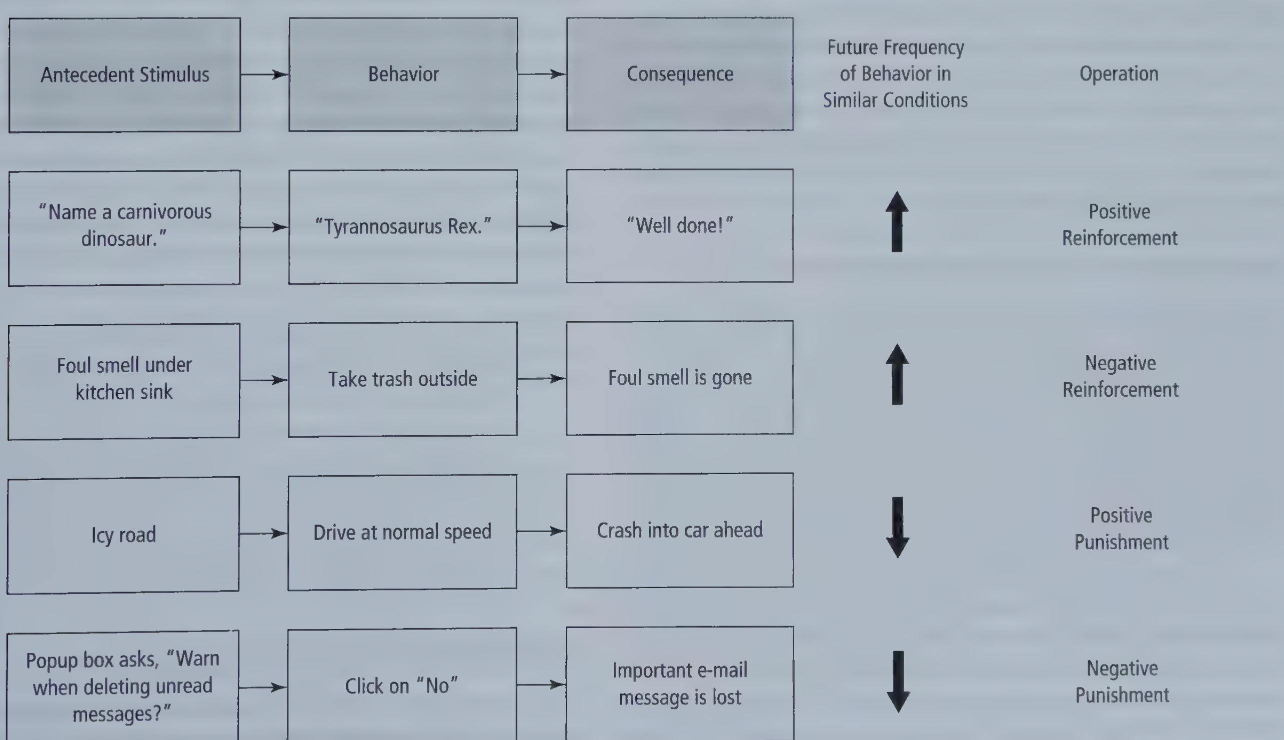


Figure 2.3 Three-term contingencies illustrating reinforcement and punishment operations.

previously, behavior is selected by the consequences that immediately follow it, irrespective of whether those consequences were produced by or depended on the behavior. This is the meaning of contingency in Skinner's (1953) statement, "So far as the organism is concerned, the only important property of the contingency is temporal" (1953, p. 85).

RECOGNIZING THE COMPLEXITY OF HUMAN BEHAVIOR

> *Behavior—human or otherwise—remains an extremely difficult subject matter.*
>
> —B. F. Skinner (1969, p. 114)

The experimental analysis of behavior has discovered a number of basic principles—statements about how behavior works as a function of environmental variables. These principles, several of which have been introduced in this chapter, have been demonstrated, verified, and replicated in thousands of experiments; they are scientific facts.[23] Tactics for changing behavior derived from these principles have been applied to a wide range of human behaviors in natural settings and analyzed in increasingly sophisticated and effective ways. An introduction to what has been learned from those applied behavior analyses makes up the majority of this book.

The systematic application of behavior analysis techniques sometimes produces behavior changes of great magnitude and speed, even for clients whose behavior had been unaffected by other forms of treatment and appeared intractable. When such a happy (but not rare) outcome occurs, the neophyte behavior analyst must resist the tendency to believe that we know more than we do about the prediction and control of human behavior. As acknowledged in Chapter 1, applied behavior analysis is a young science that has yet to achieve anything near a complete understanding and technological control of human behavior.

A major challenge facing applied behavior analysts is dealing with human behavior in applied settings where laboratory controls are impossible, impractical, or unethical. Numerous factors contributing to this difficulty coalesce around the complexity of the human repertoire, the myriad controlling variables, and individual differences.

Complexity of the Human Repertoire

Humans are capable of learning an incredible range of behaviors. Response sequences, sometimes of no apparent logical organization, contribute to the complexity of behavior (Skinner, 1953). In a response chain, effects produced by one response influence the emission of other responses. Returning a winter coat to the attic leads to rediscovering a scrapbook of old family photographs, which evokes a phone call to Aunt Helen, which sets the occasion for finding her recipe for apple pie, and so on.

Verbal behavior may be the most significant contributor to the complexity of human behavior (Donahoe & Palmer, 1994; Palmer, 1991; Skinner, 1957; see Chapter 18). Not only is a problem generated when the difference between saying and doing is not recognized, but verbal behavior itself is often a controlling variable for many other verbal and nonverbal behaviors. By following the contingency described in a verbal rule (e.g., "Pull the lever while simultaneously pushing the selection button.") we behave effectively (e.g., operate a complex vending machine in a foreign country) without direct experience with contingencies. "Rules can be learned more quickly than the behavior shaped by the contingencies they describe. . . . Rules are particularly valuable when contingencies are complex or unclear or for any other reason not very effective" (Skinner, 1974, p. 129). Skinner referred to behavior controlled by verbal statements as **rule-governed behavior**, distinguishing it from **contingency-shaped behavior** acquired by direct experience with contingencies.[24]

Operant learning does not always occur as a slow, gradual process. Sometimes new, complex repertoires appear quickly with little apparent direct conditioning (Epstein, 1991; Sidman, 1994). Murray Sidman's groundbreaking research on stimulus equivalence showed that under certain conditions learners acquire new skills and verbal relations without direct instruction on those skills and relations (cf., Critchfield, Barnes-Holmes, & Dougher, 2018). The new behaviors emerge as a product of instruction on related skills and relations. This research, which is found under various names in addition to stimulus equivalence, including generative instruction, equivalence-based instruction, and derived relational responding, is described in Chapters 19 and 20.

One type of rapid learning has been called *contingency adduction,* a process whereby a behavior that was initially selected and shaped under one set of conditions is recruited by a different set of contingencies and takes on a new function in the person's repertoire (Andronis, 1983; Layng & Andronis, 1984). Johnson and Layng (1992, 1994) described several examples of contingency adduction in which simple (component) skills (e.g., addition, subtraction, and multiplication facts, isolating and solving for *X* in a simple linear equation), when taught to fluency, combined without apparent instruction to form new complex (composite) patterns of behavior (e.g., factoring complex equations). Lessons can be designed in such a way that teaching select concepts and relations results in students acquiring concepts and relations that were not taught directly.

Intertwined lineages of different operants combine to form new complex operants (Glenn, 2004), which produce response products that in turn make possible the acquisition of behaviors beyond the spatial and mechanical restraints of anatomical structure.

> In the human case, the range of possibilities may be infinite, especially because the products of operant behavior have become increasingly complex in the context of evolving cultural practices. For example, anatomical constraints prevented operant flying from emerging in a human repertoire only until airplanes were constructed as behavioral products. Natural selection's leash has been greatly relaxed in the ontogeny of operant units. (Glenn et al., 1992, p. 1332)

Complexity of Controlling Variables

Behavior is selected by its consequences. This mega principle of operant behavior sounds deceptively (and naively) simple. However, "Like other scientific principles, its simple form masks the

complexity of the universe it describes" (Glenn, 2004, p. 134). The environment and its effects on behavior are complex.

Skinner (1957) noted that, "(1) the strength of a single response may be, and usually is, a function of more than one variable and (2) a single variable usually affects more than one response" (p. 227). Although Skinner was writing in reference to verbal behavior, multiple causes and multiple effects are characteristics of many behavior–environment relations.

Many behaviors result from multiple causes. In a phenomenon called **joint control** (Lowenkron, 1998, 2006; Palmer, 2006), two separate but interrelated forms of a person's own verbal behavior combine to acquire stimulus control of a response that would not have occurred in the absence of either. For example, a person writing on her laptop computer, while sitting in the kitchen, needs a particular book for a reference. She might say to herself, "I need *About Behaviorism*" and repeat the book's title (an echoic) several times while walking to the room containing her books. She reads the titles (another form of verbal behavior) of the books on the bookcase until she emits the response, "About Behaviorism," whose topography matches the self-echoic. The occurrence of these two sources of control mediates the selection of the desired book.[25]

Concurrent contingencies can also combine to make a behavior more or less likely to occur in a given situation. Perhaps we finally return our neighbor's weed trimmer not just because he usually invites us in for a cup of coffee, but also because returning the tool reduces the "guilt" we feel for keeping it for 2 weeks.

Concurrent contingencies often vie for control of incompatible behaviors. We cannot watch "Baseball Tonight" and study (properly) for an upcoming exam. Although not a technical term in behavior analysis, *algebraic summation* is sometimes used to describe the effect of multiple, concurrent contingencies on behavior. The behavior that is emitted is thought to be the product of the competing contingencies "canceling portions of each other out," as in an algebraic equation.

Behavioral covariation illustrates one type of multiple effect. For example, Lerman, Kelley, Vorndran, and Van Camp (2003) found that blocking the emission of one problem behavior decreased occurrences of that behavior but produced a collateral increase in other topographies of problem behaviors in the same response class. As another example of multiple effects, the presentation of an aversive stimulus may, in addition to suppressing future occurrences of the behavior it follows, elicit respondent behaviors and evoke escape and avoidance behaviors—three different effects from one event.

Emotions are often a commingling of respondent and operant behaviors. Interactions between respondent and operant conditioning further the difficulty of determining causal variables. For example Pavlovian (stimulus-reinforcer) contingencies embedded in stimulus control arrangements may limit the effectiveness of operant (response-reinforcer) contingencies (see Nevin, 2009). A study by Gorn (1982) suggests the possibility that pairing music a person likes (unconditioned stimulus for positive emotional responses) with an advertisement for a product (neutral stimulus) might condition emotional responses to the product functioning as a conditioned stimulus that elicits positive emotional responses that influence the operant behavior of purchasing the product.

All of these complex, concurrent, interrelated contingencies make it difficult for behavior analysts to identify and control relevant variables. It should not be surprising that the settings in which applied behavior analysts ply their trade are sometimes described as places where "reinforcement occurs in a noisy background" (Vollmer & Hackenberg, 2001, p. 251).

Consequently, as behavior analysts, we should recognize that meaningful behavior change might take time and many trials and errors as we work to understand the interrelationships and complexities of the controlling variables. Don Baer (1987) recognized that some of the larger problems that beset society (e.g., poverty, substance addiction, illiteracy), given our present level of technology, might be too difficult to solve at the moment. He identified three barriers to solving such complex problems:

> (a) We are not empowered to solve these bigger remaining problems, (b) we have not yet made the analysis of how to empower ourselves to try them, and (c) we have not yet made the system-analytic task analyses that will prove crucial to solving those problems when we do empower ourselves sufficiently to try them. . . . In my experience, those projects that seem arduously long are arduous because (a) I do not have a strong interim reinforcer compared to those in the existing system for status quo and must wait for opportunities when weak control may operate, even so, or (b) I do not yet have a correct task analysis of the problem and must struggle through trials and errors. By contrast (c) when I have an effective interim reinforcer and I know the correct task analysis of this problem, long problems are simply those in which the task analysis requires a series of many behavior changes, perhaps in many people, and although each of them is relatively easy and quick, the series of them requires not so much effort as time, and so it is not arduous but merely tedious. (pp. 335, 336–337)

Individual Differences

People often respond differently, sometimes drastically so, to the same set of environmental conditions. This fact is sometimes cited as evidence that principles of behavior based on environmental selection do not exist, at least not in a form that could provide the basis for a robust and reliable technology of behavior change. It is then argued that because people often respond differently to the same set of contingencies, control of behavior must come from within each person. In refuting that argument, Sidman (2013) recognizes the complex learning histories that account for much of our individual differences and points to the excitement of discovering the scientific principles that link us with one another and with the universe at large.

> Among the most powerful sources of behavioral control are our individual behavioral histories, in each of us a web of such complexity that it tempts us to give up and claim independence. . . . to claim control by an inner self. . . . Rather than needing to claim myself as prepotent, I have always found it exciting to recognize that I am at one with the rest

of the universe and to help discover the laws that relate me to everything else. (p. xvi)

As each of us experiences varying contingencies of reinforcement (and punishment), some behaviors are strengthened (selected by the contingencies) and others are weakened. This is the nature of operant conditioning, which is to say, human nature. Because no two people ever experience the world in exactly the same way, each of us arrives at a given situation with a different **history of reinforcement**. The repertoire of behaviors each person brings to any situation has been selected, shaped, and maintained by his or her unique history of reinforcement. Each human's unique repertoire defines him or her as a person. We are what we do, and we do what we have learned to do. "He begins as an organism and becomes a person or self as he acquires a repertoire of behavior" (Skinner, 1974, p. 231).

Individual differences in responding to current stimulus conditions, then, do not need to be attributed to differences in internal traits or tendencies, but to the orderly result of different histories of reinforcement. The behavior analyst must also consider people's varying sensitivities to stimuli (e.g., hearing loss, visual impairment) and differences in response mechanisms (e.g., cerebral palsy) and design program components to ensure that all participants have maximum contact with relevant contingencies (Heward, Alber-Morgan, & Konrad, 2017).

Obstacles to Changing Behavior in Applied Settings

Compounding the difficulty of tackling the complexity of human behavior in the "noisy" applied settings where people live, work, and play, applied behavior analysts are sometimes prevented from implementing an effective behavior change program due to logistical, financial, sociopolitical, legal, and/or ethical factors. Most applied behavior analysts work for agencies with limited resources, which may make the data collection required for a more complete analysis impossible. In addition, participants, parents, administrators, and even the general public may at times limit the behavior analyst's options for effective intervention (e.g., "We don't want students working for tokens"). Legal or ethical considerations may also preclude determining experimentally the controlling variables for an important behavior. Ethical considerations for behavior analysts are discussed in Chapter 31.

Each of these practical complexities combines with the behavioral and environmental complexities previously mentioned to make the applied behavior analysis of socially important behavior a challenging task. However, the task need not be overwhelming, and few tasks are as rewarding or as important for the betterment of humankind.

It is sometimes argued that the scientific study of behavior cannot account for such human creations as beauty and humor or the behavior of judging something to be beautiful. Hineline (2005), Palmer (2018), Mechner (2017), and Neuringer and Jensen (2013) discuss aesthetics and creativity from a behavior analytic point of view on topics ranging from literature aesthetics to the craftsmanship of woodworking tools.

Will knowing how such amazing behavior is learned somehow diminish the quality or enjoyment of the human experience? For example, will our increasing knowledge of the variables responsible for creative behavior lessen the feelings evoked by a powerful painting or a beautiful symphony, or reduce our appreciation of the artists who produced them? We think not, and we encourage you, as you read and study about the basic concepts introduced in this chapter and examined in more detail throughout the book, to consider Nevin's (2005) response to how a scientific account of behavior adds immeasurably to the human experience:

> At the end of *Origin of Species* (1859), Darwin invites us to contemplate a tangled bank, with its plants and its birds, its insects and its worms; to marvel at the complexity, diversity, and interdependence of its inhabitants; and to feel awe at the fact that all of it follows from the laws of reproduction, competition, and natural selection. Our delight in the tangled bank and our love for its inhabitants are not diminished by our knowledge of the laws of evolution; neither should our delight in the complex world of human activity and our love for its actors be diminished by our tentative but growing knowledge of the laws of behavior. Tony Nevin, personal communication (December 19, 2005.)

SUMMARY

Behavior

1. Behavior is the activity of living organisms.
2. Technically, behavior is "that portion of an organism's interaction with its environment that involves movement of some part of the organism" (Johnston & Pennypacker, 2009, p. 31).
3. The term *behavior* is usually used in reference to a larger set or class of responses that share certain topographical dimensions or functions.
4. *Response* refers to a specific instance of behavior.
5. *Response topography* refers to the physical shape or form of behavior.
6. A response class is a group of responses of varying topography, all of which produce the same effect on the environment.
7. *Repertoire* can refer to all of the behaviors a person can do or to a set of behaviors relevant to a particular setting or task.

Environment

8. Environment is the physical setting and circumstances in which the organism or referenced part of the organism exists.

9. Stimulus is "an energy change that affects an organism through its receptor cells" (Michael, 2004, p. 7).

10. The environment influences behavior primarily by stimulus change, not static stimulus conditions.

11. Stimulus events can be described formally (by their physical features), temporally (by when they occur), and functionally (by their effects on behavior).

12. A stimulus class is a group of stimuli that share specified common elements along formal, temporal, and/or functional dimensions.

13. Antecedent conditions or stimulus changes exist or occur prior to the behavior of interest.

14. Consequences are stimulus changes that follow a behavior of interest.

15. Stimulus changes can have one or both of two basic effects on behavior: (a) an immediate but temporary effect of increasing or decreasing the current occurrences of the behavior and/or (b) a delayed but relatively permanent effect in terms of the future occurrences of that type of behavior in the future.

Respondent Behavior

16. Respondent behavior is elicited by antecedent stimuli.

17. A reflex is a stimulus–response relation consisting of an antecedent stimulus and the respondent behavior it elicits (e.g., bright light–pupil contraction).

18. All healthy members of a given species are born with the same repertoire of unconditioned reflexes.

19. An unconditioned stimulus (e.g., food) and the respondent behavior it elicits (e.g., salivation) are called unconditioned reflexes.

20. Conditioned reflexes are the product of respondent conditioning: a stimulus–stimulus pairing procedure in which a neutral stimulus is presented with an unconditioned stimulus until the neutral stimulus becomes a conditioned stimulus that elicits the conditioned response.

21. Pairing a neutral stimulus with a conditioned stimulus can also produce a conditioned reflex—a process called higher-order (or secondary) respondent conditioning.

22. Respondent extinction occurs when a conditioned stimulus is presented repeatedly without the unconditioned stimulus until the conditioned stimulus no longer elicits the conditioned response.

Operant Behavior

23. Operant behavior is selected by its consequences.

24. Unlike respondent behavior, whose topography and basic functions are predetermined, operant behavior can take a virtually unlimited range of forms.

25. Selection of behavior by consequences operates during the lifetime of the individual organism (ontogeny) and is a conceptual parallel to Darwin's natural selection in the evolutionary history of a species (phylogeny).

26. Operant conditioning, which encompasses reinforcement and punishment, refers to the process and selective effects of consequences on behavior:
    - Consequences can affect only future behavior.
    - Consequences select response classes, not individual responses.
    - Immediate consequences have the greatest effect.
    - Consequences select any behavior that precedes them.
    - Operant conditioning occurs automatically.

27. Most stimulus changes that function as reinforcers or punishers can be described as either (a) a new stimulus added to the environment or (b) an already present stimulus removed from the environment.

28. Positive reinforcement: A response is followed by the presentation of a stimulus that results in similar responses occurring more often.

29. Negative reinforcement: A response is followed by the withdrawal of a stimulus that results in similar responses occurring more often.

30. The term *aversive stimulus* is often used to refer to stimulus conditions whose termination functions as reinforcement.

31. Extinction (withholding all reinforcement for a previously reinforced behavior) produces a decrease in response rate to the behavior's pre-reinforcement level.

32. Positive punishment: A response is followed by the presentation of a stimulus that results in similar responses occurring less often.

33. Negative punishment: A response is followed by the withdrawal of a stimulus that results in similar responses occurring less often.

34. A principle of behavior describes a functional relation between behavior and one or more of its controlling variables that has thorough generality across organisms, species, settings, and behaviors.

35. A behavior change tactic is a technologically consistent method for changing behavior that has been derived from one or more basic principles of behavior.

36. Unconditioned reinforcers and punishers function irrespective of any prior learning history.

37. Stimulus changes that function as conditioned reinforcers and punishers do so because of previous pairing with other reinforcers or punishers.

38. One important function of motivating operations is altering the current value of stimulus changes as reinforcement or punishment. For example, deprivation and satiation are

motivating operations that make food more or less effective as reinforcement.

39. A discriminated operant occurs more often under some antecedent conditions than it does under others, an outcome called stimulus control.

40. Stimulus control refers to differential rates of operant responding observed in the presence or absence of antecedent stimuli. Antecedent stimuli acquire the ability to control operant behavior by having been paired with certain consequences in the past.

41. The three-term contingency—*a*ntecedent, *b*ehavior, and *c*onsequence—is the basic unit of analysis in the analysis of operant behavior.

42. If a reinforcer (or punisher) is contingent on a particular behavior, the behavior must be emitted for the consequence to occur.

43. All applied behavior analysis procedures involve manipulation of one or more components of the three-term contingency.

Recognizing the Complexity of Human Behavior

44. Humans are capable of acquiring a huge repertoire of behaviors. Response chains and verbal behavior also make human behavior extremely complex.

45. The variables that govern human behavior are often highly complex. Many behaviors have multiple causes.

46. Individual differences in histories of reinforcement and organic impairments also make the analysis and control of human behavior difficult.

47. Applied behavior analysts are sometimes prevented from conducting an effective analysis of behavior because of practical, logistical, financial, sociopolitical, legal, and/or ethical reasons.

KEY TERMS

automatic reinforcement
automaticity of reinforcement
aversive stimulus
behavior
behavior change tactic
conditioned punisher
conditioned reflex
conditioned reinforcer
conditioned stimulus
consequence
contingency
contingency-shaped behavior
contingent
deprivation
discriminated operant
discriminative stimulus (S^D)
environment
extinction
habituation
higher-order conditioning
history of reinforcement
joint control
motivating operation
negative punishment
negative reinforcement
neutral stimulus
ontogeny
operant behavior
operant conditioning
phylogeny
positive punishment
positive reinforcement
principle of behavior
punisher
punishment
reflex
reinforcement
reinforcer
repertoire
respondent behavior
respondent conditioning
respondent extinction
response
response class
rule-governed behavior
selectionism
socially mediated contingency
stimulus
stimulus class
stimulus control
stimulus–stimulus pairing
three-term contingency
unconditioned punisher
unconditioned reinforcer
unconditioned stimulus

MULTIPLE-CHOICE QUESTIONS

1. The controlling variables of primary importance in applied behavior analysis are located in:
   a. The environment
   b. Consequences
   c. The mental states
   d. Reinforcement
   Hint: (See "Behavior")

2. Two functionally distinct types of behavior important to behavior analysis include:
   a. Learned and operant behavior
   b. Reflexive and respondent behavior
   c. Respondent and operant behavior
   d. Operant and conditioned behavior
   Hint: (See "Behavior")

3. Which of the following is considered a property of behavior amenable to measurement?
   a. Temporal Weight
   b. Temporal Locus
   c. Temporal Displacement
   d. Temporal Size
   Hint: (See "Behavior")
4. A _____ denotes a set or collection of knowledge and skills a person has learned that are relevant to particular settings or tasks.
   a. Repertoire
   b. Behavior
   c. Skill
   d. Response
   Hint: (See "Behavior")
5. Behavior is affected by stimulus changes that occur prior to and immediately after the behavior. The term _____ refers to environmental conditions or stimulus changes that exist or occur prior to the behavior of interest.
   a. Consequence
   b. Stimulus
   c. Event
   d. Antecedent
   Hint: (See "Environment")
6. A behavior that is elicited by antecedent stimuli and is "brought about" by a stimulus that precedes it is:
   a. Operant
   b. Learned
   c. Respondent
   d. New
   Hint: (See "Respondent Behavior")
7. A response is followed by a stimulus change, the effect of which is an increase in the future frequency of behavior. Which type of stimulus-change operation most likely occurred?
   a. Reinforcement
   b. Punishment
   c. Extinction
   d. Positive Punishment
   Hint: (See "Operant Behavior")
8. Water is an example of a ______ reinforcer, while money is an example of a _______ reinforcer.
   a. Conditioned, unconditioned
   b. Secondary, primary
   c. Unconditioned, conditioned
   d. Powerful, ineffective
   Hint: (See "Reinforcement")
9. Answering the door when you hear the door bell and not when it is silent is an example of behavior being under:
   a. Simultaneous prompting
   b. Equivalence
   c. Stimulus control
   d. Premack principle
   Hint: (See "Stimulus Changes That Function as Reinforcers and Punishers")
10. A behavior that occurs more frequently under some antecedent conditions than it does in others is called a(n):
    a. Stimulus control
    b. Operant behavior
    c. Discriminative stimulus
    d. Discriminated operant
    Hint: (See "Stimulus Changes That Function as Reinforcers and Punishers")
11. Food deprivation may alter the momentary effectiveness of food as a reinforcer. Food deprivation is an example of a(n):
    a. Satiation trial
    b. Motivating operation
    c. Operant behavior
    d. Experimental control
    Hint: (See "Unconditioned Reinforcement and Punishment")
12. The three-term contingency is made of these three terms:
    a. Attitude, behavior, reinforcement
    b. Antecedent, reinforcement, punishment
    c. Antecedent, behavior, consequence
    d. Antecedent, stimulus, control
    Hint: (See "The Discriminated Operant and Three-Term Contingency")
13. The term "contingent" as used in behavior analysis refers to the dependent relationship of a particular consequence on the occurrence of behavior and is also used in reference to the _________ contiguity of behavior and its consequences.
    a. Temporal
    b. Independent
    c. Dependent
    d. False
    Hint: (See "The Discriminated Operant and Three-Term Contingency")
14. _______ behavior is elicited by antecedent stimuli.
    a. Operant
    b. Temporal
    c. Respondent
    d. All
    Hint: (See "Behavior")

15. Conditioned _______ are the product of respondent conditioning.
    a. Antecedents
    b. Stimuli
    c. Punishers
    d. Reflexes
    Hint: (See "Behavior")
16. Operant behavior is selected by its:
    a. Consequences
    b. Antecedents
    c. Conditioners
    d. Respondent
    Hint: (See "Behavior")
17. Operant conditioning encompasses:
    a. Time and stimuli
    b. Antecedents and phylogeny
    c. Conditioned and unconditioned
    d. Reinforcement and punishment
    Hint: (See "Behavior")

ESSAY-TYPE QUESTIONS

1. Discuss the difference between a response and a behavior.
   Hint: (See "Behavior")
2. Discuss the importance of a functional analysis of behavior over a structural analysis of behavior.
   Hint: (See "Behavior")
3. Describe the effect of positive and negative reinforcement on subsequent behavior.
   Hint: (See "Operant Conditioning")
4. Give an example of negative reinforcement.
   Hint: (See "Operant Conditioning")
5. Discuss the difference between a principle of behavior and a behavior-change tactic.
   Hint: (See "Reinforcement")
6. Discuss the difference between conditioned and unconditioned reinforcers.
   Hint: (See "Reinforcement")
7. Describe the effect punishment contingencies have on subsequent behavior.
   Hint: (See "Selection by Consequences")
8. State an example of stimulus control.
   Hint: (See "The Discriminated Operant and Three-Term Contingency")
9. What effects do motivating operations have on reinforcers?
   Hint: (See "Unconditioned Reinforcement and Punishment")
10. Using an example from applied settings, construct a three-term contingency.
    Hint: (See "The Discriminated Operant and Three-Term Contingency")
11. Diagram an example of respondent conditioning. Identify and label the following: unconditioned stimulus, unconditioned response, neutral response, conditioned stimulus, and conditioned response.
    Hint: (See "Respondent Conditioning")
12. Classify the following behaviors as respondent or operant:
    - Talking
    - Walking
    - Baby's grasp
    - Salivation at smell of food
    - Baking a cake
    - Pupil constriction
    - Thinking about music
    - Playing the piano
    - Withdrawing hand from a fire

    Hint: (See "Respondent Conditioning")

NOTES

1. The reader should not be overwhelmed by the many technical terms and concepts in this chapter. With the exception of respondent behavior, each concept introduced in this chapter is explained in detail in subsequent chapters. This initial overview of basic concepts and principles provides background information that will facilitate understanding portions of the text that precede the more detailed explanations.
2. In the mid-1960s, Ogden Lindsley proposed what he called the *dead man test* as a way to help teachers determine whether they were targeting real behaviors for measurement and change as opposed to inanimate states such as "being quiet" (see Chapter 3). According to this test, if a dead man can do it, it isn't behavior. Getting blown over by strong wind is not behavior (a dead man would also be blown over), but moving arms and hands in front of one's face, tucking and rolling, and yelling "Whoa!" as one is being blown over are behaviors. Critchfield (2016) provides some historical background on the dead man test and a thoughtful critique of its usefulness as a guide for measurement.
3. Most behavior analysts use the word *behavior* both as a mass noun to refer to the subject matter of the field in general or a certain type or class of behavior (e.g., operant behavior, study behavior) and as a count noun to refer to specific instances (e.g., two aggressive behaviors). The word *behavior* is often implied and unnecessary to state. We agree with Friman's (2004) recommendation that in most cases, "If the object of our interest is hitting and spitting,

let's just say 'hitting' and 'spitting.' Subsequently, when we are gathering our thoughts with a collective term, we can call them behaviors" (p. 105).

4. Respondent conditioning and the operant principles mentioned here are introduced later in this chapter.
5. Gray (1979) provides an excellent and interesting description of Pavlov's research. "Although the experimental preparations for studying Pavlovian conditioning have expanded over the years, Pavlov's work with dogs describes many of the key empirical phenomena and theoretical processes that modern researchers continue to pursue. . . . The study of basic associative learning processes underlying Pavlovian conditioning has led to numerous insights into experimental design, how learning occurs, and how basic processes may lay the foundation for many putative higher forms of learning" (K. M. Lattal, 2013, p. 283).
6. *Unconditioned stimulus* and *conditioned stimulus* are the most commonly used terms to denote the stimulus component of respondent relations. However, because the terms ambiguously refer to both the immediate evocative (eliciting) effect of the stimulus change and its somewhat permanent and delayed function-altering effect (the conditioning effect on other stimuli), Michael (1995) recommended that the terms *unconditioned elicitor* (UE) and *conditioned elicitor* (CE) be used when referring to the evocative function of these variables.
7. To learn more about Pavlovian conditioning, see the detailed descriptions and numerous examples in Gottlieb and Begej (2014), K. M. Lattal (2013), and Williams (2014).
8. The verb *emit* is used in conjunction with operant behavior. Its use fits well with the definition of operant behavior, allowing reference to the consequences of behavior as the major controlling variables. The verb *elicit* is inappropriate to use with operant behavior because it implies that an antecedent stimulus has primary control of the behavior. Domjan (2016) argues that the terms *elicited* and *emitted* are no longer useful or accurate for distinguishing classically conditioned (respondent) behavior from operant behavior.
9. To learn more about the role of selectionism in Darwinian evolutionary biology and Skinnerian behaviorism, see Leão, Laurenti, and Haydu (2016); Moxley (2004); and Reese (1994).
10. Unless noted otherwise, the term *behavior* will refer to operant behavior throughout the remainder of the text.
11. Skinner (1966) used rate of responding as the fundamental datum for his research. To strengthen an operant is to make it more frequent. However, rate is not the only measurable and malleable dimension of behavior. As we will see in Chapters 3 and 4, sometimes the duration, latency, magnitude, and/or topography of behavior changes are of practical importance.
12. *Automaticity of reinforcement* and *automatic reinforcement* are different concepts. **Automatic reinforcement** refers to the hypothesis that a behavior produces its own reinforcement, as in the sensory stimulation from scratching an itch (Vaughn & Michael, 1982). In practice, when an experimental analysis fails to identify the reinforcers maintaining a problem behavior, some form of *automatic reinforcement* is assumed. Automatic reinforcement is described in Chapters 11 and 27.
13. Prominent behavior analysts disagree on whether delayed consequences can function as reinforcement (Bradley & Poling, 2010). Chapters 11, 29, and 30 discuss the use and effects of delayed consequences in behavior change programs.
14. Malott and Shane (2014) refer to these two operations as "stimulus addition" and "stimulus subtraction."
15. For examples of inaccurate representations of behaviorism and behavior analysis and their implications for training practitioners and serving clients, see Cooke (1984); Heward (2005); Heward and Cooper (1992); Morris (2009); Schlinger (2018); and Todd and Morris (1983, 1992).
16. Foxx (1982) introduced the terms *Type I punishment* and *Type II punishment* for positive punishment and negative punishment, respectively. Some behavior analysts and practitioners continue to use Foxx's terminology. Malott and Shane (2014) refer to negative punishment as a *penalty contingency*.
17. Michael (1975) and Baron and Galizio (2005) present cogent arguments that positive and negative reinforcement are examples of the same fundamental operant relation. This issue is discussed further in Chapter 12.
18. Some authors use the modifiers *primary* or *unlearned* to identify *unconditioned reinforcers* and *unconditioned punishers*.
19. In addition to using *aversive stimulus* as a synonym for a *negative reinforcer*, Skinner (1953) used the term to refer to stimuli whose onset or presentation functions as punishment, a practice continued by many behavior analysts (e.g., Alberto & Troutman, 2013; Malott & Shane, 2014; Miltenberger, 2016). The term *aversive stimulus* (and *aversive control* when speaking of behavior change techniques involving such stimuli) is used widely in the behavior analysis literature to refer to one or more of three different behavioral functions: An aversive stimulus may be (a) a negative reinforcer if its termination increases behavior, (b) a punisher if its presentation decreases behavior, and/or (c) a motivating operation if its presentation increases the current frequency of behaviors that have terminated it in the past (see Chapter 16). When speaking or writing technically, behavior analysts must be careful that their use of omnibus terms such as *aversive* does not imply unintended functions (Michael, 1995).
20. Some authors use the modifiers *secondary or learned* to identify *conditioned reinforcers* and *conditioned punishers*.
21. Contingency diagrams, such as those shown in Figure 2.3, are an effective way to illustrate temporal and functional relations between behavior and environmental events. Examples of other types of contingency diagrams and suggestions for using them to teach and learn about behavior analysis can be found in Goldwater and Acker (1995); Malott and Shane (2014); Mattaini (1995); and Toogood (2012). State notation is another means for visualizing complex contingency relations and experimental procedures (Mechner, 1959; Michael & Shafer, 1995).
22. The phrase *to make reinforcement contingent* describes the behavior of the researcher or practitioner: delivering the reinforcer only after the target behavior has occurred.
23. Like all scientific findings, these facts are subject to revision and even replacement should future research reveal more useful data.
24. In *shaping*, a new behavior is gradually acquired through reinforcement of successive approximations to a final form (see Chapter 22). Because the natural environment seldom (if ever) truly shapes new behavior, *contingency-selected behavior* may be a better term than *contingency-shaped behavior*.
25. This example was adapted from Sidener (2006). More examples of joint control are presented in Chapter 18, Verbal Behavior.

PART 2

Selecting, Defining, and Measuring Behavior

Because applied behavior analysts must achieve and document behavior changes that improve people's quality of life, the careful selection and systematic measurement of behavior form the operational foundation of applied behavior analysis. Chapter 3 describes the role of assessment in behavior analysis, principal assessment methods used by behavior analysts, how to identify and assess the social significance of potential target behaviors, how to prioritize target behaviors, and how to define the selected behaviors to enable accurate and reliable measurement, and discusses options for setting criteria for behavior change. Chapter 4 explains the role of measurement in applied behavior analysis, defines measurable dimensions of behavior, describes procedures and tools for measuring behavior, and suggests guidelines for selecting a measurement system. Chapter 5 identifies indicators of trustworthy measurement; elucidates common threats to the validity, accuracy, and reliability of behavioral measurement; recommends strategies for combating those threats; and describes methods for assessing the accuracy and quality of behavioral measurement.

CHAPTER 3

Selecting and Defining Target Behaviors

LEARNING OBJECTIVES

- Understand and explain the role of assessment in applied behavior analysis.
- Identify the ethical and professional standards of professional applied behavior analysis.
- State target intervention outcomes in observable and measurable terms.
- Describe the four major methods for obtaining assessment information.
- Explain the importance of social validity in regard to selecting target behavior.
- Describe procedures for assessing the social significance of potential target behaviors.
- Discuss criteria for prioritizing target behaviors.
- Define behavior in observable and measurable terms.
- Explain the process for setting criteria for behavior change.

Applied behavior analysis is concerned with producing predictable and replicable improvements in behavior. However, not just any behavior will do: Applied behavior analysts improve *socially significant behaviors* that have immediate and long-lasting effects for the person and for those who interact with that person. Such behaviors include language, social, motor, and academic skills.

An important preliminary and universally accepted first step involves choosing the *right* behaviors to target for assessment, measurement, and change (Lerman, Iwata, & Hanley, 2013). Lerman et al. (2013) state: "Specific target behaviors are usually selected because they will improve a person's quality of life in the short and the long term by allowing the person access to new reinforcers and additional reinforcing contexts" (p. 86).

This chapter describes the role of assessment in applied behavior analysis, three principal behavioral assessment methods, how to assess and prioritize the social significance of potential target behaviors, and how to apply criteria for defining target behaviors.

ROLE OF ASSESSMENT IN APPLIED BEHAVIOR ANALYSIS

Assessment has long been considered the first of four phases in any systematic model of instruction that includes: assessment, planning, implementation, and evaluation (Stephens, 1976). A comprehensive behavioral assessment informs planning, provides guidance for implementation, and assists with evaluation.

Definition and Purpose of Behavioral Assessment

Traditional psychological and educational assessments typically involve a series of norm-referenced, criterion-referenced, or curriculum-based assessments supplemented by observations, anecdotal reports, and historical data to determine the learner's strengths and weaknesses within cognitive, academic, social, and/or psychomotor domains.

Conversely, **behavioral assessment** includes indirect and direct procedures such as interviews, checklists, and tests to identify and define the specific target behavior. In addition to identifying behavior(s) to change, comprehensive behavioral assessment can uncover functional relations between variables; it provides context on the resources, assets, significant others, competing contingencies, maintenance and generalization factors, and reinforcers (or punishers) that can be combined to improve the efficiency of an intervention (Brown, McDonnell, & Snell, 2016).[1]

Wacker, Berg, Harding, and Cooper-Brown (2011) indicated that the principal goal of behavioral assessment is "to identify environmental variables that are related to increases or decreases in target behavior" (p. 165). Linehan (1977) offered a succinct description of the purpose of behavioral assessment: "To figure out what the client's problem is and how to change it for the better" (p. 31). Implicit in the definitions of both Wacker et al. and Linehan is the concept that behavioral assessment is more than an exercise in describing and classifying behavioral abilities and deficiencies. In short, behavioral assessment goes beyond trying to obtain a psychometric score, grade equivalent data, or rating measure, as worthy as such findings might be for other purposes. Behavioral assessment discovers the *function* that behavior serves in the person's environment (e.g., obtaining social attention, escaping or avoiding a task). A well-constructed and thorough assessment identifies the key variables controlling the behavior, and guides the practitioner to manipulate these variables to the benefit of the learner. Consequently, subsequent interventions can be targeted and have a much better chance of success. As Bourret, Vollmer, and Rapp (2004) pointed out: "The critical test of . . . assessment is the degree to which it differentially indicates an effective teaching strategy" (p. 140).

Phases of Behavioral Assessment

Hawkins's (1979) conceptualization of behavioral assessment embraces five phases: (a) screening, (b) defining and quantifying problems and establishing outcome criteria, (c) pinpointing the target behavior(s), (d) monitoring progress, and (e) following up. Although the five phases form a general chronological sequence, overlap often occurs. Part III of this book, "Evaluating and Analyzing Behavior Change," describes the monitoring and follow-up stages of assessment. This chapter is concerned primarily with the preintervention functions of assessment, the selection and definition of a **target behavior**—the specific behavior selected for change.

To serve competently, applied behavior analysts must know what constitutes socially important behavior, have the technical skills to use appropriate assessment methods and instruments, analyze the data, and be able to match assessment data with an intervention strategy. For instance, the remedial reading specialist must understand the critical behaviors of a competent reader; be able to determine the learner's progress across a continuum of recognition, decoding, and comprehension skills; and deliver appropriate and effective instruction. In short, the analyst must be knowledgeable about the full scope and sequence of the target behavior.

Preassessment Considerations

Before conducting an informal or formal behavioral assessment for the purpose of pinpointing a target behavior, the analyst must address two fundamental questions. First, who has the authority, permission, resources, and skills to complete an assessment and intervene with the client? If a practitioner does not have authority or permission, then the analyst's role in assessment and intervention is restricted. For example, suppose that a behavior analyst is standing in a checkout line where a nearby parent is attempting to manage an extremely disruptive child. Does the behavior analyst have the authority to conduct an on-the-spot assessment of the problem or to suggest an intervention to the parent? No. However, if the same episode occurred during an *in situ* "coaching session" with the parent who had requested assistance beforehand, the analyst would be able to offer an assessment and recommend an intervention.

The second question asks: What records, resources, or data currently exist that shed light on what has been done in the past to identify, treat, and evaluate the target behavior? Minimally, the behavior analyst should examine and review relevant medical, educational, or historical data. Clearly, all medical reasons for suspected unusual behavior should be ruled out. A child with serious, but as yet undetected, eye (amblyopia), ear (otitis media), or prenatal conditions (fetal alcohol syndrome) may exhibit behavior problems, which, if treated solely by behavioral measures, will miss the mark. Pretreating these causal factors medically allows for better focus in later educational and/or behavioral interventions.

Likewise, the practitioner should conduct an analysis of prior educational and behavioral interventions. Through examining narrative reports, analyzing data-based graphs or charts, and interviewing caretakers, a skilled analyst can form a picture about the nature of the target behavior, its developmental scope and sequence, and the reinforcers and punishers that appear to sustain the behavior. Further, a judgment can be made about treatment integrity with prior interventions. It's possible that a past intervention (e.g., extinction) was employed to reduce calling out in class, but it was abandoned prematurely because the caretakers misinterpreted the *gradual* decrease in call outs as failure. Alternatively, misbehaviors that were addressed successfully in a training environment (e.g., special education classroom, developmental center) might surface in an untrained setting (home, community setting) because of a past failure to program for generalization.

In effect, applied behavior analysts must not only recognize the role of assessment in the four-stage assessment-planning-implementation-evaluation continuum, but also recognize all the personal, social, and environmental factors that affect behavior. Only then can they use their behavior analytic skills to assess and ultimately change behavior.[2] Figure 3.1 outlines the four cornerstones of professional assessment.

Figure 3.1 Four cornerstones of professional assessment.

1. *Recognize professional limitations.*	• Secure authority to assess the problem or suggest an intervention. • Do not test beyond professional training or license capability. • Decline to assess under poor conditions (e.g., the learner is ill, the environment is distracting, the analyst is rushed for time).
2. *Assume responsibility for your work.*	• Be careful, be thorough, be conservative with interpretations. • Recognize the assumptions for referral and testing (i.e., change of placement, generate intervention strategies).
3. *Maintain confidentiality.*	• Hold assessment information in confidence. Obtain a signed and dated Release of Information (ROI). • Ensure that assessment results are shared only with those who are eligible to receive them. • Seek guidance in situations where confidentiality is in question (e.g., life-threatening situations).
4. *Adhere to professional, administration, and ethical standards.*	• Use technically reliable and valid measures and procedures to assess. • Follow appropriate protocols. • Examine and analyze all relevant medical, educational, and historical data. • Recommend the strongest, but least intrusive interventions.

ASSESSMENT METHODS USED BY BEHAVIOR ANALYSTS

Behavior analysts use a variety of assessment methods to identify which behavior(s) to target for change. A comprehensive behavioral assessment often includes both indirect and direct methods.[3]

Indirect Assessment

Interviews, checklists, and rating scales are *indirect assessment approaches* because the data obtained from these methods are derived from recollections, reconstructions, and/or subjective, ordinal-scale ratings of events. Interviews, checklists, and rating scales can be conducted with the client and/or with people who come into regular contact with the individual (e.g., teachers, parents, care providers).

Indirect assessments include open- and closed-ended formats (Fryling & Baires, 2016). *Open-ended indirect assessments* encourage the informant to comment freely and without restriction on the number, rate, intensity, and duration of target behaviors and stimuli and conditions that surround the occurrence of these behaviors (*How often does Mario engage in self-injurious behavior?*). The informant's narrative across a range of questions often leads to follow-up questions that provide more context and substance on the possible variables controlling the behavior (*Does Mario's self-injurious behavior occur more often when he is at home or at school?*).

Closed-ended indirect assessments require the informant to use a Likert scale to rate a series of questions, with a total summary score being generated that provides direction on possible variables controlling the behavior (e.g., *Rate the occurrences of Mario's self-injurious behavior using this scale:* 5 = *frequently*; 3 = *sometimes*; 1 = *rarely*). Figure 3.2 compares the relative strengths and limitations of closed- and open-ended indirect assessments.

The following sections describe the roles of the client, significant others, and checklists and rating scales with respect to conducting indirect assessments.

Interviewing the Client

A behavioral interview is often a first and important step in identifying potential target behaviors, which can be verified, or rejected, by subsequent direct observation or empirical analysis. A behavioral interview differs from a traditional interview by the type of questions asked and the level of information sought. Behavior analysts rely primarily on *what* and *when* questions, which tend to focus on the environmental conditions that exist before, during, and after a behavioral episode. Identifying environmental events that correlate with the behavior provides valuable information for formulating hypotheses about the controlling function of these variables and for planning interventions. Hypothesis generation leads to experimental manipulation and the discovery of functional relations (see Chapter 26).

Figure 3.3 provides examples of the appropriate kinds of *what* and *when* questions. This sequence of questions was developed by a behavioral consultant in response to a teacher who wanted to reduce the frequency of her own negative reactions to acting out and disruptive students. Similar questions could be generated to address situations in homes or community settings (Sugai & Tindal, 1993).

Interviewing Significant Others

Sometimes either the behavior analyst cannot interview the client personally or the analyst needs information from others who are important in the client's life (e.g., parents, teachers, coworkers). In such cases, the analyst will interview one or more of these significant others. When asked to describe a behavioral problem or deficit, significant others often begin with general terms or labels that do not identify specific behaviors to change and often imply causal factors thought to be intrinsic to the client (e.g., fearful, aggressive, unmotivated, lazy, withdrawn). By asking structured questions, the behavior analyst helps significant others describe the problem in terms of specific behaviors, environmental conditions, and events associated with those behaviors. For example, when interviewing parents, the following questions help to target instances of "noncompliant" and "immature" behaviors.

- What does Derek do that leads you to characterize his behavior as immature or noncompliant?
- During what time of day does Derek seem most immature (or noncompliant)?

Figure 3.2 Strengths and limitations of closed- versus open-ended indirect assessments.

	Closed-ended indirect assessments	Open-ended indirect assessments
Strengths	-Assures information about common controlling variables with some populations. -Quick and easy to administer. -Few particular skills are required to administer and interpret.	-Information about a wide range of contextual variables may be obtained. -Opportunity to develop rapport with the informant and to listen to his or her experiences.
Limitations	-Only asks questions about pre-determined variables. -May suggest a non-function (i.e., false-positives). -May be misused by individuals without training in ABA. -Little opportunity to develop rapport with informants.	-Takes longer than closed-ended indirect assessments. -May include information that is not relevant to the function. -Interpretation requires behavior analytic skills. -Requires clinical interviewing skills to obtain information and develop rapport.

- Are there certain situations or places where Derek is noncompliant or acts immature? If so, where, and what does he do?
- How many different ways does Derek act immature (or noncompliant)?
- What's the most frequent noncompliant behavior that Derek has?
- How do you and other family members respond when Derek exhibits these behaviors?
- If Derek were to be more mature and independent as you would like, what would he do differently than he does now?

Figure 3.4 is a form with which parents, caregivers, and significant others can identify possible target behaviors.

Figure 3.3 Sample behavioral interview questions.

Problem Identification Interview Form

Reason for referral: The teacher requested help reducing her negative attention to acting out and disruptive students who were yelling and noncompliant.

1. In your own words, can you define the problem behaviors that prompted your request?
2. Are there any other teacher-based behaviors that concern you at this time?
3. When you engage in negative teacher attention (i.e., when you attend to yelling or noncompliant behavior), what usually happens *immediately before* the negative teacher attention occurs?
4. What usually happens *after* the negative teacher attention occurs?
5. What are the students' reactions when you yell or attend to their noncompliant behavior?
6. What behaviors would the students need to perform so that you would be less likely to attend to them in a negative way?
7. Have you tried other interventions? What has been their effect?

Figure 3.4 Form for generating possible target behaviors.

The 5 + 5 Behavior List

Child's name: ______________________________

Person completing this list: ______________________________

Listmaker's relationship to child: ______________________________

5 good things _______ does now	5 things I'd like to see _______ learn to do more (or less) often
1. ______________________	1. ______________________
2. ______________________	2. ______________________
3. ______________________	3. ______________________
4. ______________________	4. ______________________
5. ______________________	5. ______________________

Directions: Begin by listing in the left-hand column 5 desirable behaviors your child (or student) does regularly now; things that you want him or her to continue doing. Next, list in the right-hand column 5 behaviors you would like to see your child do more often (things that your child does sometimes but should do with more regularity) and/or undesirable behaviors that you want him or her to do less often (or not at all). You may list more than 5 behaviors in either column, but try to identify at least 5 in each.

In addition to seeking help from significant others to identify target behaviors and possible controlling variables, the behavior analyst can sometimes use the interview to determine the extent to which significant others are willing and able to help implement a subsequent intervention. Without the assistance of parents, siblings, teacher aides, and staff, it is difficult for behavior change programs to succeed, or be sustained, with any degree of treatment integrity after the initial intervention.

As an outgrowth of being interviewed, some clients, or their significant others, may be asked to complete a questionnaire or so-called *needs assessment survey*. Questionnaires and needs assessment surveys have been developed in multiple human services areas to refine or extend the interview process (Altschuld & Witkin, 2000). Results from such assessments can be useful in selecting and defining target behaviors further, or in determining possible interventions. For example, a client seeking behavioral treatment to quit smoking might respond to a questionnaire or survey by indicating the number of cigarettes smoked each day and the conditions under which smoking occurs (e.g., morning coffee break, after dinner, stuck in a traffic jam). These client-collected data may shed light on antecedent conditions correlated with the target behavior.

Checklists and Rating Scales

Behavior checklists and rating scales can be used alone or in combination with interviews to identify potential target behaviors (Kelley, LaRue, Roane & Gadaire, 2011). A **behavior checklist** provides a description of specific behaviors and the conditions under which each behavior occurs (i.e., antecedent and consequence events that may affect the frequency, intensity, or duration of behaviors). Situation- or program-specific checklists can be created to assess one particular behavior (e.g., tooth brushing) or a specific skill area (e.g., social skills).

A rating scale using a Likert system (e.g., a range of numbers from 1 to 5 or 1 to 10) attempts to more precisely describe and quantify the target behavior of interest by using an ordinal scale. When a numerically based scale is employed, sometimes a total score is generated to provide further guidance on the significance of the target behavior. The *Functional Assessment Checklist for Teachers and Staff* (March et al., 2000) combines checklist and rating scale components to help specify the target behavior and the conditions under which it occurs (see Figure 3.5). Still, whenever an ordinal scale is used, and data interpretation is involved, behavior analysts are advised to heed Merbitz, Morris, and Grip's (1989) sage advice:

> Ordinal level scales are widely used in rehabilitation. Such scales provide a rank order characterization of phenomena when the direction of an underlying dimension (such as consciousness) can be described and units of measurement cannot. Unfortunately, the information produced by these scales is difficult to interpret and easy to misuse. (p. 308)

Direct Assessment

Tests and direct observations are *direct assessment* approaches because the results from these procedures provide measurable and validated information about the learner's behavior. Standardized tests and observations are two principal methods of direct assessment.

Standardized Tests

Literally thousands of standardized tests have been developed to assess behavior (cf. Carlson, Geisinger, & Jonson, 2017). Each time a *standardized test* is administered, the same questions and tasks are presented using specified procedures and the same scoring criteria are applied. Some standardized tests yield norm-referenced scores, meaning that during the test's development it was administered to a large sample of people selected to represent the population for whom the test is designed, and scoring tables and charts are produced that can later be referenced to plot, chart, or calculate results. Scores are charted on a normalized curve. See Box 3.1: "Primer on Norm-referenced Test Scores".

The majority of standardized tests on the market, however, are not conducive to behavioral assessment because the results cannot be translated directly into target behaviors for instruction or treatment. For example, results from standardized tests commonly used in the schools, such as the *Iowa Tests of Basic Skills* (Hoover, Dunbar, & Frisbie, 2007), the *Peabody Individual Achievement Test–R/NU* (Markwardt, 2005), and the *Wide Range Achievement Test—5* (WRAT-5) (Wilkinson & Robertson, 2017), might indicate that a fourth-grader is performing at the third-grade level in mathematics and at the first-grade level in reading. Such information is useful in determining how the student performs with these subjects compared to students in general, but it indicates neither the specific math nor reading skills the student has mastered. Further, it does not provide sufficient direction with which to launch a remedial, intervention, or enrichment program.

Finally, and from a practical perspective, behavior analysts may not be able to administer a given standardized test because of licensing requirements. For instance, only a licensed psychologist can administer some types of intelligence and personality inventories.

Tests are most useful as behavioral assessment devices when they provide a specific measure of the person's performance of the behaviors of interest. Practitioners often use criterion-referenced and curriculum-based assessments to help target behaviors to change.

Criterion-referenced Assessment

A criterion-referenced assessment (CRA) measures a child's skill performance across recognized and commonly accepted "developmental milestones." For instance, most 12-month-old toddlers can sit with trunk stability, stand independently, walk a few steps, reach across their midline, and roll a ball with an adult. In scoring CRAs, teachers might indicate a plus (+), minus (−), or (E) for each associated skill measured to signify that the skill was present, not present, or emerging, respectively.

Curriculum-based Assessment

A curriculum-based assessment (CBA), by contrast, is especially useful because the data that are obtained bear uniquely on the

Figure 3.5 The Functional Assessment Checklist for Teachers and Staff.

Functional Assessment Checklist for Teachers and Staff (FACTS-Part A)

Step 1 Student/Grade: ______________________ Date: ______________________
Interviewer: ______________________ Respondent(s): ______________________

Step 2 **Student Profile:** Please identify at least three strengths or contributions the student brings to school.
____

Step 3 **Problem Behavior(s): Identify problem behaviors**

____ Tardy	____ Fight/physical Aggression	____ Disruptive	____ Theft
____ Unresponsive	____ Inappropriate Language	____ Insubordination	____ Vandalism
____ Withdrawn	____ Verbal Harassment	____ Work not done	____ Other ________
	____ Verbally Inappropriate	____ Self-injury	

Describe problem behavior: ____

Step 4 **Identifying Routines: Where, when and with whom problem behaviors are most likely**

Schedule (Times)	**Activity**	**Likelihood of Problem Behavior**						**Specific Problem Behavior**
	Before School	Low 1	2	3	4	5	High 6	
	Math	1	2	3	4	5	6	
	Transition	1	2	3	4	5	6	
	Language Arts	1	2	3	4	5	6	
	Recess	1	2	3	4	5	6	
	Reading	1	2	3	4	5	6	
	Lunch	1	2	3	4	5	6	
	Science	1	2	3	4	5	6	
	Transition	1	2	3	4	5	6	
	Block Studies	1	2	3	4	5	6	
	Art	1	2	3	4	5	6	

Step 5 **Select 1–3 routines for further assessment: Select routines based on (a) similarity of activities (conditions) with ratings of 4, 5 or 6 and (b) similarity of problem behavior(s). Complete the FACTS-Part B for each routine identified.**

Functional Assessment Checklist for Teachers & Staff (FACTS-Part B)

Step 1 Student/Grade: ______________________ Date: ______________________
Interviewer: ______________________ Respondent(s): ______________________

Step 2 **Routine/Activities/Context:** Which routine (only one) from the FACTS-Part A is assessed?

Routine/Activities/Context	**Problem Behavior(s)**

(continued)

Figure 3.5 (continued)

Step 3 **Provide more detail about the problem behavior(s):**

What does the problem behavior(s) look like?

How often does the problem behavior(s) occur?

How long does the problem behavior(s) last when it does occur?

What is the intensity/level of danger of the problem behavior(s)?

Step 4 **What are the events that predict when the problem behavior(s) will occur? (Predictors)**

Related Issues (setting events)		**Environmental Features**	
____ illness	Other: ____________	____ reprimand/correction	____ structured activity
____ drug use	____________	____ physical demands	____ unstructured time
____ negative social	____________	____ socially isolated	____ tasks too boring
____ conflict at home	____________	____ with peers	____ activity too long
____ academic failure	____________	____ Other	____ tasks too difficult
		____________	

Step 5 **What consequences appear most likely to maintain the problem behavior(s)?**

Things that are Obtained		**Things Avoided or Escaped From**	
____ adult attention	Other: ____________	____ hard tasks	Other: ____________
____ peer attention	____________	____ reprimands	____________
____ preferred activity	____________	____ peer negatives	____________
____ money/things	____________	____ physical effort	____________
		____ adult attention	____________

SUMMARY OF BEHAVIOR

Step 6 **Identify the summary that will be used to build a plan of behavior support.**

Setting Events & Predictors	**Problem Behavior(s)**	**Maintaining Consequence(s)**

Step 7 **How confident are you that the Summary of Behavior is accurate?**

Not very confident					**Very confident**
1	2	3	4	5	6

Step 8 **What current efforts have been used to control the problem behavior?**

Strategies for preventing problem behavior		**Strategies for responding to problem behavior**	
____ schedule change	None: ____ Other ____	____ reprimand	None: ____ Other ____
____ seating change	____________	____ office referral	____________
____ curriculum change	____________	____ detention	____________

BOX 3.1

Primer on Norm-Referenced Test Scores

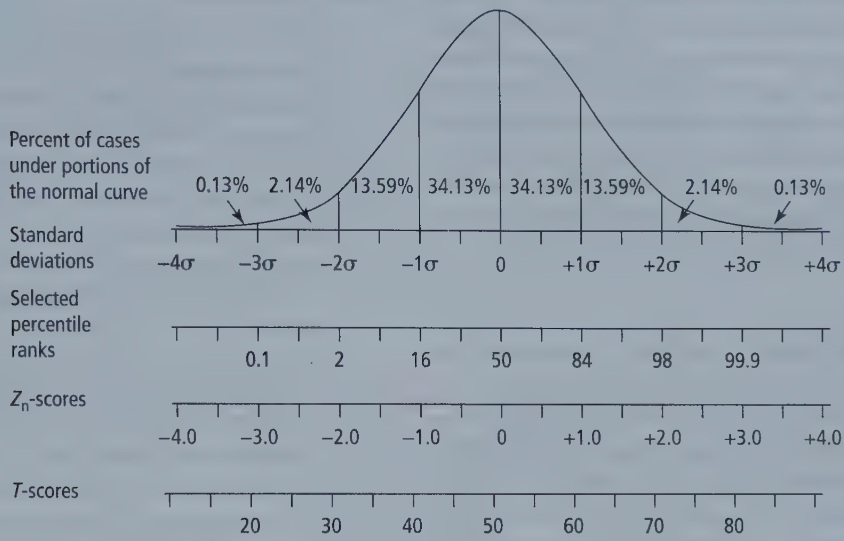


Figure A Relationship among percentile scores, *z*-scores, and *t*-scores in a normal distribution.

From *Educational Assessment of Students*, 8th ed., p. 391, by S. M. Brookhart and A. J. Nitko, 2019, Upper Saddle River, NJ: Pearson. Reprinted by permission.

Norm-referenced tests compare an individual's performance with a comparable norm group. These tests provide a number of metrics that practitioners use to make decisions. Three widely used and reported norm-referenced scores are percentiles, *z*-scores, and *t*-scores. Figure A shows the relationships of these and other norm-referenced scores within a normal distribution

Standardized tests yield a number of metrics that practitioners use to make decisions. Plotted beneath this "normal curve" are key measures that can be used to compare an individual's performance with that of a comparable norm group. Test scores of most interest are percentile ranks, *z*-scores, and *t*-scores.

Percentile Ranks

A percentile rank aligns a score with a standard deviation. An average score is the 50th percentile, located in the middle of the distribution. A score of 84 means that the tester's score was equal to or better than the scores of 84% of the individuals who took the standardized test and equates with 1 standard deviation above the mean.

Z-scores

A *z*-score represents the number of standard deviations that a score appears below, at, or above the mean of the population who completed the assessment. A *z*-score of +1.0 equates to 1 standard deviation, and also the 84th percentile.

T-scores

A *t*-score is a standard score with an average of 50 and a standard deviation of 10. In our highlighted example, a *t*-score of 50, located in the middle of the distribution, equates to a percentile score of 50 (read the percentile scale above).

Two other derived scores of interest are the grade equivalent score and the age equivalent (not shown on the normal curve).

Grade Equivalent Score

A grade equivalent score—say, a GES of 5.5—means that the learner's raw score (the number of total items answered correctly on the test) equaled the average number raw score by learners midway through fifth grade in the norm group. This score does not provide information on the specific skills the individual possesses or lacks.

Age Equivalent Score

An age equivalent score—say, 4.6—means that the learner's raw score (the number of total items answered correctly on the test) equaled the average number of raw score points earned by learners midway through age 4 in the norm group. This score does not provide information on the specific skills the individual possesses or lacks.

To learn more about standardized testing and interpreting norm-referenced scores, see assessment texts such as those by Brookhart and Nitko (2019) and Salvia, Ysseldyke, and Bolt (2013).

daily or weekly tasks the learner performs in response to planned lessons (i.e., the curriculum) (Overton, 2006). A learner taking a spelling test at the end of the week based on a 5-day list of words, or a middle school student taking a science quiz on a chapter from the text, serves as a prime example of curriculum-based assessments. CBA data can be collected on a formative basis (e.g., to note error patterns or relative progress) or on a summative schedule (e.g., yearly) across motor (gross and fine), language (receptive and expressive), self-help, or cognition areas. Teachers would record aggregate scores or, depending on the learner's developmental level, mark associated tasks within the skill area as occurring independently, semi-independently, or not at all.

Direct Observations

Direct and repeated observations of the client's behavior in the natural environment are useful for determining behaviors to assess and ultimately to select as a target behavior. Practitioners, however, should be advised that observations could produce limited and skewed information. For instance, if data are collected only during morning sessions, at home, or with a specific caretaker, the "target" behavior may be masked in the afternoon, at school, or with different caretakers. Expanding the direct observations to include a wider range of settings, persons, or behaviors is likely to improve the overall quality of direct assessment, but until the manipulation of variables occurs, it will fall short of being a complete analysis (Thompson & Borrero, 2011).

One basic form of direct continuous observation, first described by Bijou, Peterson, and Ault (1968), is called ABC, or **anecdotal observation**, recording. With **ABC recording**, the observer records a descriptive, temporally sequenced account of all behaviors of interest and the antecedent conditions and consequences for those behaviors as those events occur in the client's natural environment (Cooper, 1981). This technique produces behavioral assessment data that can be used to identify potential target behaviors.[4]

Anecdotal observation yields an overall description of a client's behavior patterns. This detailed record of the client's behavior within its natural context provides accountability to the individual and to others involved in the behavior change plan and is extremely helpful in designing interventions.

Accurately describing behavioral episodes as they occur in real time is aided by using a form to record relevant antecedents, behaviors, and consequences in temporal sequence. For example, Lo (2003) used the form shown in Figure 3.6 to record anecdotal observations of a fourth-grade special education student whose teacher had complained that the boy's frequent talk-outs and out-of-seat behavior were impairing his learning and often disrupted the entire class. ABC observations can also be recorded on a checklist of specific antecedents, behaviors, and consequent events individually created for the client based on information from interviews and/or initial observations (see Figure 27.3).

ABC recording requires the observer to commit full attention to the person being observed for upwards of 20 to 30 minutes. Hence, a classroom teacher, for example, could not easily use this assessment procedure while engaging in other activities, such as managing a reading group, demonstrating a mathematics problem on the chalkboard, or conducting a classwide peer tutoring session. Following are some additional guidelines and suggestions for conducting anecdotal direct observations:

- Record everything the client does and says and everything that happens to the client.
- Use homemade shorthand or abbreviations to make recording more efficient, but be sure the notes can be and are accurately expanded immediately after the observation session.
- Record only actions that are seen or heard, not interpretations of those actions.
- Record the temporal sequence of each response of interest by writing down what happened just before and just after it.
- Record the estimated duration of each instance of the client's behavior. Mark the beginning and ending time of each behavioral episode.
- Be aware that continuous anecdotal observation is often an obtrusive recording method. Most people behave differently when they see someone with a pencil and clipboard staring at them. Knowing this, observers should be as unobtrusive as possible (e.g., stay a reasonable distance away from the subject, avoid eye contact).
- Carry out the observations over a period of several days so that the novelty of having someone observe the client will lessen and the repeated observations can produce a valid picture of day-to-day behavior.

Ecological Assessment

Behavior analysts understand that human behavior is a function of multiple events and that many events have multiple effects on behavior (cf. Michael, 1995). An ecological approach to assessment recognizes the complex interrelationships between environment and behavior. In an **ecological assessment** a great deal of information is gathered about the person and the various environments in which that person lives and works. Among the many factors that can affect a person's behavior are physiological conditions, physical aspects of the environment (e.g., lighting, seating arrangements, noise level), interactions with others, home environment, and past reinforcement history. Each of these factors represents a potential area for assessment.

Although a thorough ecological assessment will provide a tremendous amount of descriptive data, the basic purpose of assessment—to identify the most pressing behavior problem and possible ways to alleviate it—should not be forgotten. It is easy to go overboard with the ecological approach, gathering far more information than necessary. Ecological assessment can be costly in terms of professional and client time, and it may raise ethical and perhaps legal questions regarding confidentiality (Koocher & Keith-Spiegel, 1998). Ultimately, good judgment must be used in determining how much assessment information is necessary. Writing about the role of ecological assessment for special education teachers, Heron and Heward (1988) suggested that

> The key to using an ecological assessment is to know when to use it. . . . Educators must strive to become keen discriminators of: (1) situations in which a

Figure 3.6 Example of an anecdotal ABC recording form.

Student: Student 4 Date: 3/10/03 Setting: SED resource room (math period)

Observer: Experimenter Starting time: 2:40 P.M. Ending time: 3:00 P.M.

Time	Antecedents (A)	Behavior (B)	Consequences (C)
2:40	T tells students to work quietly on their math worksheets	Walks around the room and looks at other students	T says, "Everyone is working, but you. I don't need to tell you what to do."
	✓	Sits down, makes funny noises with mouth	A female peer says: "Would you please be quiet?"
	✓	Says to the peer, "What? Me?" Stops making noises	Peer continues to work
2:41	Math worksheet	Sits in his seat and works quietly	No one pays attention to him
	Math worksheet	Pounds on desk with hands	SED aide asks him to stop
2:45	Math worksheet	Makes vocal noises	Ignored by others
	Math worksheet	Yells-out T's name three times and walks to her with his worksheet	T helps him with the questions
2:47	Everyone is working quietly	Gets up and leaves his seat	T asks him to sit down and work
	✓	Sits down and works	Ignored by others
	Everyone is working quietly	Gets up and talks to a peer	T asks him to sit down and work
	✓	Comes back to his seat and works	Ignored by others
2:55	Math worksheet, no one is attending to him	Hand grabs a male peer and asks him to help on the worksheet	Peer refuses
	✓	Asks another male peer to help him	Peer helps him
2:58	✓	Tells T he's finished the work and it's his turn to work on computer	T asks him to turn in his work and tells him that it's not his turn on the computer
	✓	Whines about why it's not his turn	T explains to him that other students are still working on the computer and he needs to find a book to read
	✓	Stands behind a peer who is playing computer and watches him play the game	Ignored by T

Adapted from *Functional Assessment and Individualized Intervention Plans: Increasing the Behavior Adjustment of Urban Learners in General and Special Education Settings* (p. 317) by Y. Lo. Unpublished doctoral dissertation. Columbus, OH: The Ohio State University. Used by permission.

planned intervention has the potential for affecting student behaviors other than the behavior of concern; and (2) situations in which an intervention, estimated to be effective if the target behavior were viewed in isolation, may be ineffective because other ecological variables come into play. (p. 231)

Reactive Effects of Assessment

Reactivity refers to the effects that an assessment has on the behavior being assessed. Reactivity is most likely when observation is *obtrusive*—that is, the person being observed is aware of the observer's presence and purpose (Kazdin, 2010). Studies have demonstrated that the presence of observers in

applied settings can influence a subject's behavior (Mercatoris & Craighead, 1974; White, 1977). Perhaps the most obtrusive assessment procedures are those that require the participant to monitor and record her own behavior. Research on self-monitoring shows that the procedure commonly affects the behavior under assessment (Kirby, Fowler, & Baer, 1991).[5]

Research suggests, however, that even when the presence of an observer alters the behavior of the person being observed, the reactive effects are temporary (e.g., Kazdin, 1982). Nevertheless, behavior analysts mitigate reactivity when they make observations as unobtrusive as possible, repeat observations until apparent reactive effects subside, and take possible reactive effects into account when interpreting the results of observations.

Choosing an Assessment Method

In addition to selecting assessment methods and tools for identifying target behaviors, behavior analysts conduct assessments to guide the design of effective interventions. These require the experimental manipulation of variables and are described elsewhere in this text: preference and reinforcer assessment (Chapter 11), punisher assessment (Chapter 14), and functional behavior analysis (Chapter 27).

Given that multiple possibilities exist across assessment methods, how might a behavior analyst proceed when trying to weigh the cost-benefit of delaying an intervention in favor of gathering more "assessment" data versus launching an intervention prematurely? Kelley and colleagues (2011) provide a guideline:

> Ultimately, best practice is likely to include an assessment package that contains a structured interview, at least one rating scale, direct observation, and experimental manipulation of environmental variables. A combination of these assessment methods is likely to produce . . . a desirable outcome. (p. 189)

We restate a key point: Assuming the assessment tools have reasonable reliability and validity, and are conducted according to professional standards, the most important element of the assessment process is not so much the gathering of the data as it is the *analysis* of that data. If the analysis should prove not to produce a desirable outcome as Kelley et al. (2011) project, practitioners should not be dismayed. Instead, they should try again "based on the original failure and previous ineffective analysis. . . . Failure is always informative in the logic of behavior analysis" (Baer, 2005, p. 8).

ASSESSING THE SOCIAL SIGNIFICANCE OF POTENTIAL TARGET BEHAVIORS

In the past, few questions were asked when a teacher, therapist, or other human services professional determined that a client's behavior should be assessed. For a well-trained behavior analyst, this is not the case. Because behavior analysts possess an effective technology to change behavior, accountability must be served. Both the goals and the rationale supporting assessment and intervention programs must be open to critical examination by the consumers (i.e., clients and their families) and by others who may be affected (i.e., society). In selecting target behaviors, practitioners should first consider *whose* behavior is being assessed—and changed—and why.

Target behaviors should not be selected solely for the primary benefit of others (e.g., "Be still, be quiet, be docile," in Winett & Winkler, 1972), to simply maintain the status quo (Budd & Baer, 1976; Holland, 1978), or because they pique the interest of someone in a position to change the behaviors. For instance, assume that a newly hired director of an early childhood developmental daycare center announces an annual goal that all the children learn to talk, when in fact, it would be a more appropriate option to assess (and later intervene) on a fuller range of communication options (e.g., signing, picture exchange communication). Selecting "speaking" as a target goal for all children stands in stark contrast to the more fundamental and laudable goal of providing specialized services to meet the individual needs of the children.

Admittedly, judgments about which behaviors to change are difficult to make. Still, practitioners are not without direction when choosing target behaviors. Numerous authors have suggested guidelines and criteria for selecting target behaviors (e.g., Bailey & Lessen, 1984; Bosch & Fuqua, 2001; Hawkins, 1984; Komaki, 1998; Rosales-Ruiz & Baer, 1997). Overall, these guidelines revolve around this central question: To what extent will the proposed behavior change improve the person's life experience? A key answer to this question rests squarely on the concept of habilitation.

Habilitation Defined

Hawkins (1984) suggested that the potential meaningfulness of any behavior change should be considered within the context of **habilitation**, which he defined as "the degree to which the person's repertoire maximizes short- and long-term reinforcers for that individual and for others, and minimizes short- and long-term punishers." (p. 284)

Hawkins (1986) cited several advantages of the definition in that it (a) is conceptually familiar to behavior analysts, (b) defines treatment using measurable outcomes, (c) is applicable to a wide range of habilitative activities, (d) deals with individual and societal needs nonjudgmentally, (e) treats adjustment along a non-deficit-driven, adaptive continuum, and (f) is culturally and situationally relative.

It is difficult to make a priori decisions about how a particular behavior change will contribute to a person's overall habilitation (adjustment, competence). We simply do not know beforehand how useful or functional a given behavior change will prove to be in the future (Baer, 1981, 1982), even when its short-term utility can be predicted. Applied behavior analysts, however, must place the highest importance on the selection of target behaviors that are truly useful and habilitative (Hawkins, 1991). In effect, if a potential target behavior meets the habilitation standard, then the individual is much more likely to (a) acquire additional reinforcers in the future and (b) avoid potential punishers.

From ethical and pragmatic perspectives, any target behavior must benefit the person either directly or indirectly. Examining prospective target behaviors according to the 10 key questions described in the following sections should help clarify their relative social significance and habilitative value.

Figure 3.7 Worksheet for evaluating the social significance of potential target behaviors.

Client's/Student's Name: ________________________ Date: ____________

Person Completing Worksheet: ________________________

Rater's Relationship to Client/Student: ________________________

Behavior: ________________________

Considerations	Assessment	Rationale/Comments
Is this behavior likely to produce reinforcement in the client's natural environment after intervention ends?	Yes No Not Sure	
Is this behavior a necessary prerequisite for a more complex and functional skill?	Yes No Not Sure	
Will this behavior increase the client's access to environments where other important behaviors can be acquired and used?	Yes No Not Sure	
Will changing this behavior predispose others to interact with the client in a more appropriate and supportive manner?	Yes No Not Sure	
Is this behavior a pivotal behavior or behavioral cusp?	Yes No Not Sure	
Is this an age-appropriate behavior?	Yes No Not Sure	
If this behavior is to be reduced or eliminated from the client's repertoire, has an adaptive and functional behavior been selected to replace it?	Yes No Not Sure	
Does this behavior represent the actual problem/goal, or is it only indirectly related?	Yes No Not Sure	
Is this "just talk," or is it the real behavior of interest?	Yes No Not Sure	
Is the goal itself is not a specific behavior (e.g., losing 20 lb), will this behavior help achieve it?	Yes No Not Sure	

Summary notes/comments: ________________________

Figure 3.7 consolidates these questions in a worksheet that can be used to evaluate the social significance of target behaviors.

Will This Behavior Produce Reinforcement in the Client's Natural Environment After Treatment Ends?

To determine whether a particular target behavior is functional for the client, the behavior analyst, significant others, and, whenever possible, the client should ask whether the proposed behavior change will be reinforced in the person's daily life. Ayllon and Azrin (1968) called this the **relevance of behavior rule**; it means that a target behavior should be selected only when it can be determined that the behavior is likely to produce reinforcement in the person's natural environment. The likelihood that a new behavior will result in reinforcement after the behavior change program is terminated is the primary determinant of whether the new behavior will be maintained, thereby having the possibility of long-term benefits for that person.

Judging whether occurrences of the target behavior will be reinforced in the absence of intervention can also help to clarify whether the proposed behavior change is primarily for the individual's benefit or for someone else's. For instance, despite parental wishes or pressure, it would be of little value to try to teach math skills to a student with severe developmental disabilities with pervasive deficits in communication and social skills. Teaching communication skills that would enable the student to have more effective interactions in her *current* environment should take precedence over skills that she might be able to use in the *future* (e.g., making change in the grocery store). Sometimes target behaviors are selected appropriately not because of their direct benefit to the person, but because of an important indirect benefit. Indirect benefits can occur in several different ways, as described by the questions that follow.

Is This Behavior a Necessary Prerequisite for a Useful Skill?

Some behaviors that, in and of themselves, are not important are targeted for instruction because they are necessary prerequisites to learning other functional behaviors. For example, advances in reading research have demonstrated that teaching phonemic awareness skills has positive effects on the acquisition of reading skills (e.g., initial, medial, and final isolated and blended sound combinations) (National Reading Panel, 2000).[6]

Will This Behavior Increase the Client's Access to Environments Where Other Important Behaviors Can Be Learned and Used?

Hawkins (1986) described the targeting of "access behaviors" as a means of producing indirect benefits to clients. For example, students in special education are taught to complete their assignments, interact politely and appropriately with the teacher and their peers, and follow teacher instructions because these behaviors not only will help their personal progress in their current setting but also will likely increase their acceptance and adaptation into a general education classroom.

Will Changing This Behavior Predispose Others to Interact with the Client in a More Appropriate and Supportive Manner?

Another type of indirect benefit occurs when a behavior change is of primary interest to a significant other in the person's life. The behavior change may enable the significant other to behave in a manner more beneficial to the person. For example, suppose a teacher wants the parents of his students to implement a home-based instruction program, believing that the students' language skills would improve considerably if their parents spent just 10 minutes per night playing a vocabulary game with them. In meeting with one student's parents, however, the teacher realizes that although the parents are also concerned about poor language skills, they have other and, in their opinion, more pressing needs—the parents want their child to clean her room and help with the dinner dishes. Even though the teacher believes that straightening up a bedroom and washing dishes are not as important to the child's ultimate welfare as language development, these tasks may be important target behaviors if a sloppy room and a sink full of dirty dishes impede positive parent–child interactions (including playing the teacher's vocabulary-building games). In this case, the daily chores might be selected as the target behaviors for the direct, immediate benefit of the parents, with the expectation that the parents will be more likely to help their daughter with school-related activities if they are happier with her because she straightens her bedroom and helps with the dishes.

Is This Behavior a Behavioral Cusp or a Pivotal Behavior?

Behavior analysts often use a building block method to develop repertoires for clients. For example, in teaching a complex skill (e.g., two-digit multiplication), simpler and more easily attainable skills are taught first (e.g., addition, regrouping, single-digit multiplication), or with shoe tying, crossing laces, making bows, and tying knots are taught systematically. As skill elements are mastered, they combine into increasingly complex behaviors. At any point along this developmental skill continuum when the person performs the skill to criterion, reinforcement follows, and the practitioner makes the determination to advance to the next skill level. As systematic and methodical as this approach has proven to be, behavior analysts are researching ways to improve the efficiency of developing new behavior (see Chapters 19 through 23). Choosing target behaviors that are behavioral cusps and pivotal behaviors may improve this effort.

Behavioral Cusps

A **behavioral cusp** occurs when a learner performs a new behavior that sets the occasion to access reinforcers that otherwise would not have been available. Stated in different terms, a behavioral cusp can be claimed when what has been learned exposes the individual to an array of contingencies that go beyond the reinforcers that would have been delivered had only the target behavior, or a series of behaviors, occurred.

Rosales-Ruiz and Baer (1997) defined a behavioral cusp as:

> [A] behavior that has consequences beyond the change itself, some of which may be considered important. . . . What makes a behavior change a cusp is that it exposes the individual's repertoire to new environments, especially new reinforcers and punishers, new contingencies, new responses, new stimulus controls, and new communities of maintaining or destructive contingencies. When some or all of those events happen, the individual's repertoire expands; it encounters a differentially selective maintenance of the new as well as some old repertoires, and perhaps that leads to some further cusps. (p. 534)

Rosales-Ruiz and Baer (1997) cited crawling, reading, and generalized imitation as examples of behavioral cusp behaviors because the occurrence of such behaviors "*suddenly open* the child's world to new contingencies that will develop many new, important behaviors" (p. 535, italics added for emphasis).

Cusps differ from component or prerequisite behaviors in that cusps ultimately occasion new behaviors in novel environments that bring the learner under different social, sensory, or personal stimulus control (Lerman et al., 2013; Twyman, 2011). For an infant, specific arm, head, leg, or positional movements would be component behaviors for crawling. Crawling is the cusp because it enables the infant to contact new stimuli as sources of reinforcement (e.g., toys, parents), which, in turn, open a host of contingencies that can shape and select other adaptive behaviors. These novel contingencies become available for the learner *and* the significant others in the learner's community.

Rosales-Ruiz and Baer (1997) distinguish between a learner performing a series of tasks in a task analysis and receiving the associated reinforcer for completing each step, and a behavioral cusp that exposes the learner to a host of new contingencies of reinforcement altogether. In plain language, a cusp might be akin to tree branches shooting off the trunk, and providing the possibility of still more offshoots from that "cusp" branch.

Rosales-Ruiz & Baer (1997) sum up the role of identifying cusps as target behaviors by stating:

> The importance of cusps is judged by (a) the extent of the behavior changes they systematically enable, (b) whether they systematically expose behavior to new cusps, and (c) the audience's view of whether these changes are important for the organism, which in turn is often controlled by societal norms and expectations of what behaviors should develop in children and when that development should happen. (p. 537)

By identifying and assessing target behaviors based on their cusp potential, practitioners provide "added value" to their programming. Bosch and Hixson (2004) proposed a series of actions behavior analysts can take and questions to determine behavioral cusps (see Figure 3.8). If the answer is "yes" to the majority of these questions in the right-hand column, a good case can be made that the behavior is a cusp.

Finally, Twyman (2011) sheds much-needed light on how present or future technologies can assist practitioners with developing and refining behavioral cusps. Whether it be through websites, smartphones, iPad/tablet applications, social media, simulations, virtual learning applications, or access to on-line instruction, behavior analysts can further explore programming cusp behaviors to enhance the learning of those they serve.

Pivotal Behavior

A **pivotal behavior** is a behavior that, once learned, produces corresponding modifications or covariations in other adaptive untrained behaviors. For instance, Koegel, Carter, and Koegel (2003) indicated that teaching children with autism to "self-initiate" (e.g., approach others) may be a pivotal behavior. Choice making, self-management, and functional communication training may also be considered pivotal behaviors in that once these behaviors are learned, they are likely to produce adaptive variations in untrained settings.

Koegel and Koegel (2018) and their colleagues have examined pivotal behavior assessment and treatment approaches across a wide range of areas (e.g., social skills, communication, academic performance disruptive behavior) (Koegel, Bradshaw, Ashbaugh, & Koegel, 2014; Koegel, Koegel, Green-Hopkins,

Figure 3.8 Actions and questions to determine behavioral cusps.

Information Needed	Behavioral Cusp Determination
List the direct reinforcers for that target behavior: Social, automatic, conditioned, unconditioned.	Will the behavior contact new reinforcers?
List and describe the environments that the new target behavior will access.	Will the behavior allow access to new selective environments?
List school personnel or familial beliefs and expectations that may strengthen or weaken the value of the behavior change.	Does the behavior meet the demands of the social community of which the person is a member?
List the audience reactions, such as smiles, laughter, etc.	Is the behavior a reinforcer for the audience?
List the inappropriate behaviors and their severity, including frequency, duration, strength, etc. and the effects on the environment.	Is the behavior interfering with or replacing inappropriate behaviors?
List the behaviors/repertoires subsequently affected.	Is the behavior facilitating subsequent learning by being either a prerequisite or a component of more complex responses?
Approximate number of people physically or emotionally affected over time.	Is the behavior impacting a large number of people?
Give a dollar amount by estimating the cost of physical/property damage, health care, lawsuits, etc. over time.	Is the failure to establish this behavior costly?

From "The Final Piece to a Complete Science of Behavior: Behavior Development and Behavioral Cusps," by S. Bosch and M. D. Hixson, *The Behavior Analyst Today*, 5(2), 2004, pp. 244–254.

& Barnes, 2010; Koegel, Koegel, & Schreibman, 1991). The "longitudinal outcome data from children with autism suggest that the presence of pivotal initiation behaviors may be a prognostic indicator of more favorable long-term outcomes and therefore may be 'pivotal' in that they appear to result in widespread positive changes in a number of areas" (Koegel, Carter, & Koegel, 2003, p. 134). Improvement in self-initiations may be critical for the emergence of untrained response classes, such as asking questions and increased production and diversity of talking.

Assessing and targeting pivotal behaviors can be advantageous for the practitioner and the client. From the practitioner's perspective, it might be possible to assess and then train pivotal behaviors within relatively few sessions that would later be emitted in untrained settings or across untrained responses (Koegel et al., 2003). From the client's perspective, learning a pivotal behavior would shorten intervention, provide the person with a new repertoire with which to interact with his environment, improve the efficiency of learning, and increase the chances of coming into contact with reinforcers. As Koegel and colleagues concluded: "The use of procedures that teach the child with disabilities to evoke language learning opportunities in the natural environment may be particularly useful for speech and language specialists or other special educators who desire ongoing learning outside of language teaching sessions" (p. 143).

Is This an Age-Appropriate Behavior?

A number of years ago it was common to see adults with developmental disabilities being taught behaviors that an adult without disabilities would seldom, if ever, do. It was thought—perhaps as a by-product of the concept of mental age—that a 35-year-old woman with the verbal skills of a 10-year-old should play with dolls. Not only is the selection of such target behaviors demeaning, but their occurrence reduces the probability that other people in the person's environment will set the occasion for and reinforce more desirable, adaptive behaviors, leading to a more normal and rewarding life.

The principle of **normalization** refers to the use of progressively more typical environments, expectations, and procedures "to establish and/or maintain personal behaviors which are as culturally normal as possible" (Wolfensberger, 1972, p. 28). Normalization is not a single technique, but a philosophical position that holds the goal of achieving the greatest possible physical and social integration of people with disabilities into the mainstream of society.

In addition to the philosophical and ethical reasons for selecting age- and setting-appropriate target behaviors, it should be re-emphasized that adaptive, independent, and social behaviors that come into contact with reinforcement are more likely to be maintained than are behaviors that do not. For example, instruction in leisure-time skills such as sports, hobbies, and music-related activities would be more functional for a 17-year-old boy than teaching him to play with toy trucks and building blocks. An adolescent with those behaviors—even in an adapted way—has a better chance of interacting in a typical fashion with his peer group, which may help to ensure the maintenance of his newly learned skills and provide opportunities for learning other adaptive behaviors.

If the Proposed Target Behavior Is to Be Reduced or Eliminated, What Adaptive Behavior Will Replace It?

A practitioner should never plan to reduce or eliminate a behavior from a person's repertoire without first (a) determining an adaptive behavior that will take its place (also known as the *fair pair rule*) and (b) designing the intervention plan to ensure that the replacement behavior is learned. Teachers and other human services professionals should be in the business of building positive, adaptive repertoires, not merely reacting to and eliminating behaviors they find troublesome (Brown et al., 2016). Even though a child's maladaptive behaviors may be exceedingly annoying to others, or even damaging physically, those undesirable responses have proven functional for the child. That is, the maladaptive behavior has worked for the child in the past by producing reinforcers and/or helping the child avoid or escape punishers. A program that only denies that avenue of reinforcement is a nonconstructive approach. It does not teach adaptive behaviors to replace the inappropriate behavior.

Some of the most effective and recommended methods for eliminating unwanted behavior focus primarily on the development of desirable replacement behaviors. Goldiamond (1997) recommended that a "constructional" approach—as opposed to an eliminative approach—be used for the analysis of and intervention into behavioral problems. Under the constructional approach the "solution to problems is the construction of repertoires (or their reinstatement or transfer to new situations) rather than the elimination of repertoires" (Goldiamond, 1974, p. 14). Finally, using the "fair pair" rule helps to ensure that maladaptive behaviors are replaced with incompatible, adaptive behaviors.

If a strong case cannot be made for specific, positive replacement behaviors, then a compelling case has not been made for eliminating the undesirable target behavior. The classroom teacher, for example, who wants a behavior change program to maintain students staying in their seats during reading period must go beyond the simple notion that "they need to be in their seats to do the work." The teacher must select materials and design contingencies that facilitate that goal and motivate the students to accomplish their work.

Does This Behavior Represent the Actual Problem or Goal, or Is It Only Indirectly Related?

An all-too-common error in education is teaching a related behavior, not the behavior of interest. Numerous behavior change programs have been designed to increase on-task behaviors when the primary objective should have been to increase production or work output. On-task behaviors are chosen because people who are productive also tend to be on task. However, as *on task* is typically defined, it is possible for a student to be on task (e.g., in her seat, quiet, and oriented toward or handling academic materials) yet produce little or no work.

Targeting needed prerequisite skills should not be confused with selecting target behaviors that do not directly represent or fulfill the primary reasons for the behavior analysis effort. Prerequisite skills are not taught as terminal behaviors for their own sake, but as necessary elements of the desired terminal behavior. Related, but indirect, behaviors are not necessary to perform the true objective of the program, nor are they really

intended outcomes of the program by themselves. In attempting to detect indirectness, behavior analysts should ask two questions: Is this behavior a necessary prerequisite to the intended terminal behavior? Is this behavior what the instructional program is really all about? If either question can be answered affirmatively, the behavior is eligible for target behavior status.

Is This Just Talk, or Is It the Real Behavior of Interest?

A characteristic of nonbehavioral therapies is that they rely heavily on what people *say* about what they do and why they do it. The client's verbal behavior is considered important because it is believed to reflect the client's inner state and the mental processes that govern the client's behavior. Therefore, getting a person to talk differently about himself (e.g., in a more healthful, positive, and less self-effacing way) is viewed as a significant step in solving the person's problem. This change in attitude is considered by some to be the primary goal of therapy.

Behavior analysts, in contrast, distinguish between what people say and what they do (Skinner, 1953). Knowing and doing are not the same. Getting someone to understand his maladaptive behavior by being able to talk logically about it does not necessarily mean that his behavior will change in more constructive directions. The gambler may know that compulsive betting is ruining his life and that his losses would cease if he simply stopped placing bets. He may even be able to verbalize these facts to a therapist and state convincingly that he will not gamble in the future. Still, he may continue to bet.

Because verbal behavior can be descriptive of what people do, it is sometimes confused with the performance itself. A teacher at a school for juvenile offenders introduced a new math program that included instructional games, group drills, timed tests, and self-graphing. The students responded with negative comments: "This is stupid," "Man, I'm not writin' down what I do," "I'm not even going to try on these tests." If the teacher had attended only to the students' talk about the program, it would probably have been discarded on the first day. But the teacher was aware that negative comments about school and work were expected in the peer group of adolescent delinquents and that her students' negative remarks had enabled them in the past to avoid tasks they thought they would not enjoy. Consequently, the teacher ignored the negative comments and attended to and rewarded her students for accuracy and rate of math computation when they participated in the program. In 1 week, the negative talk had virtually ceased, and the students' math production was at an all-time high.

There are, of course, situations in which the behavior of interest *is* what the client says. Helping a person reduce the number of self-effacing comments he makes and increase the frequency of positive self-descriptions is an example of a program in which talk should be the target behavior—not because the self-effacing comments are indicative of a poor self-concept, but because the client's verbal behavior is the problem.

In every case, a determination must be made of exactly which behavior is the desired functional outcome of the program: Is it a skill or motor performance, or is it verbal behavior? In some instances, doing and talking behaviors might be important. A trainee applying for a lawn mower repair position may be more likely to get a job if he describes verbally how he would fix a cranky starter on a mower. However, once hired, he will hold his job if he is skilled and efficient in repairing mowers. It is highly unlikely that a person will last long on the job if he talks about how he would fix a lawn mower but not be able to do so. Target behaviors must be functional.

What if the Goal of the Behavior Change Program Is Not a Behavior?

Some of the important changes people want to make in their lives are not behaviors, but are instead the result or product of certain other behaviors. Weight loss is an example. On the surface, it appears that target behavior selection is obvious and straightforward—losing weight. The number of pounds can be measured accurately; but weight, or, more precisely, losing weight, is not a behavior. Losing weight is not a specific response that can be defined and performed; it is the product or result of other behaviors—notably reduced food consumption and/or increased exercise. Eating and exercise are definable, observable behaviors that can be measured in precise units.

Some otherwise well-designed weight loss programs have not been successful because behavior change contingencies were placed on the goal—reduced weight—and not on the behaviors necessary to produce the goal. Target behaviors in a weight loss program should be measures of food consumption and exercise level, with intervention strategies designed to address those behaviors (e.g., De Luca & Holborn, 1992; McGuire, Wing, Klem, & Hill, 1999). Weight should be measured and charted during a weight loss program, not because it is the target behavior of interest, but because weight loss shows the positive effects of increased exercise or decreased food consumption.

There are numerous other examples of important goals that are not behaviors, but are the end products of other behaviors. Earning higher grades, for example, is a goal accomplished by practicing skills longer, complying with feedback on practiced performance, joining groups that perform even higher than the learner and imitating that performance, and so forth. Behavior analysts help clients achieve their goals by selecting target behaviors that are the most directly and functionally related to those goals.

Some goals expressed by and for clients are not the direct product of a specific target behavior, but are broader, more general goals: To be more successful, to have more friends, to be creative, to learn good sportsmanship, to develop an improved self-concept. Clearly, none of these goals are defined by specific behaviors, and all are more complex in terms of their behavioral components than losing weight or earning a higher grade. Goals such as being successful represent a class of related behaviors or a general pattern of responding. They are labels that are used to describe people who behave in certain ways. Selecting target behaviors that will help clients or students attain these kinds of goals is even more difficult than their complexity suggests because the goals themselves often mean different things to different people. Being a success entails a wide variety of behaviors. One person may view success in terms of income and job title. For another, success means job satisfaction and ample leisure time. An important role of the behavior analyst during assessment and target behavior identification is to help the client select and define personal behaviors, the sum of which will result in a better quality of life for the client and others evaluating her repertoire in the intended fashion.

PRIORITIZING TARGET BEHAVIORS

Once a "pool" of eligible target behaviors has been identified, decisions must be made about their relative priority. Sometimes the information obtained from behavioral assessment points to one particular aspect of the person's repertoire in need of improvement more so than another. More often, though, assessment reveals a constellation of related, and sometimes not-so-related, behaviors in need of change. Direct observations, along with a behavioral interview and needs assessment, may produce a long list of important behaviors to change. When more than one eligible target behavior remains after careful evaluation of the considerations described in the previous section, the question becomes, Which behavior should be changed first? Judging each potential target behavior in light of the following nine questions may help determine which behavior deserves attention first, and the relative order in which the remaining behaviors will be addressed.

1. *Does this behavior pose any danger to the client or to others?* Behaviors that cause harm or pose a serious threat to the client's or to others' personal safety or health must receive first priority.
2. *How many opportunities will the person have to use this new behavior?* or *How often does this problem behavior occur?* A student who consistently writes reversed letters presents more of a problem than does a child who reverses letters only occasionally. If the choice is between first teaching a prevocational student to pack his lunch or to learn how to plan his 2-week vacation each year, the former skill takes precedence because the employee-to-be may need to pack his lunch every workday.
3. *How long-standing is the problem or skill deficit?* A chronic behavior problem (e.g., bullying) or skill deficit (e.g., lack of social interaction skills) should take precedence over problems that appear sporadically or that have just recently surfaced.
4. *Will changing this behavior produce higher rates of reinforcement for the person?* If all other considerations are equal, a behavior that results in higher, sustained levels of reinforcement should take precedence over a behavior that produces little additional reinforcement for the client.
5. *What will be the relative importance of this target behavior to future skill development and independent functioning?* Each target behavior should be judged in terms of its relation (i.e., prerequisite or supportive) to other critical behaviors needed for optimal learning and development and maximum levels of independent functioning in the future.
6. *Will changing this behavior reduce negative or unwanted attention from others?* Some behaviors are not maladaptive because of anything inherent in the behavior itself, but because of the unnecessary problems the behavior causes the client. Some people with developmental and motoric disabilities may have difficulty at mealtimes with using utensils and napkins appropriately, thus reducing opportunities for positive interaction in public. Granted, public education and awareness are warranted as well, but it would be naive not to consider the negative effects of public reaction. Also, not teaching more appropriate mealtime skills may be a disservice to the person. Idiosyncratic public displays or mannerisms may be high-priority target behaviors if their modification is likely to provide access to more normalized settings or important learning environments.
7. *Will this new behavior produce reinforcement for significant others?* Even though a person's behavior should seldom, if ever, be changed simply for the convenience of others or for maintaining the status quo, neither should the effect of a person's behavior change on the significant others in his life be overlooked. This question is answered best by the significant others themselves because people not directly involved in the person's life would often have

   > no idea how rewarding it is to see your 19-year-old with severe intellectual disabilities acquire the skill of toilet flushing on command or pointing to food when she wants a second helping. I suspect that the average taxpayer would not consider it "meaningful" to him or her for Karrie to acquire such skills. And, although we cannot readily say how much Karrie's being able to flush the toilet enhances her personal reinforcement/punishment ratio, I can testify that it enhances mine as a parent. (Hawkins, 1984, p. 285)

8. *How likely is success in changing this target behavior?* Some behaviors are more difficult to change than others. At least four sources of information help assess the level of difficulty or, more precisely, predict the ease or degree of success in changing a particular behavior. First, what does the literature say about attempts to change this behavior? Many of the target behaviors that confront applied behavior analysts have been studied. Practitioners should stay abreast of published research reports in their areas of application. Not only is such knowledge likely to improve the selection of proven and efficient techniques for behavior change, but also it may help to predict the level of difficulty or chance of success.

   Second, how experienced is the practitioner? The practitioner's own competencies and experiences with the target behavior in question should be considered. A teacher who has worked successfully with acting-out, aggressive children may have an array of effective behavior management strategies to employ. However, that same teacher may believe that he is less able to improve significant deficits in a student's written language skills.

   Third, to what extent can important variables in the client's environment be controlled effectively? Whether a certain behavior *can* be changed is not the question. In an applied setting, however, identifying and then consistently manipulating the controlling variables for a given target behavior will determine whether the behavior *will* be changed.

   Fourth, are the resources available to implement and maintain the intervention at a level of fidelity and intensity that is long enough to achieve the desired outcomes? No matter how expertly designed a treatment plan, implementing it without the personnel and other resources needed to carry out the intervention properly is likely to yield disappointing results.
9. *How much will it cost to change this behavior?* Cost should be considered before implementing any systematic behavior change program. However, a cost–benefit analysis of several potential target behaviors does not mean that if

a teaching program is expensive, it should not be implemented. Major courts have ruled that the lack of public funds may not be used as an excuse for not providing an appropriate education to all children regardless of the severity of their disability (cf., Yell & Drasgow, 2000). The cost of a behavior change program cannot be determined by simply adding aggregate expenditures for equipment, materials, transportation, staff salaries, and the like. Consideration should also be given to how much of the client's time the behavior change program will demand. If, for example, teaching a fine motor skill to a child with severe disabilities consumes so much of the child's day that there is little time for her to learn communication, leisure, or self-help skills, the fine motor skill objective may be too costly.

Developing and Using a Target Behavior Ranking Matrix

Assigning a numerical rating to each of a list of potential target behaviors produces a priority ranking of those behaviors. One such ranking matrix is shown in Figure 3.9; it is an adaptation of a system described by Dardig and Heward (1981) for prioritizing

Figure 3.9 Worksheet for prioritizing potential target behaviors.

Client's/Student's Name: ______________________ Date: __________

Person Completing Worksheet: ______________________

Rater's Relationship to Client/Student: ______________________

Directions: Use the key below to rank each potential target behavior by the extent to which it meets or fulfills each prioritizations criteria. Add each team member's ranking of each potential target behavior. The behavior(s) with the highest total scores would presumably be the highest priority for intervention. Other criteria relevant to a particular program or individual's situation can be added, and the criteria can be differentially weighted.

Key: 0 = No/Never; 1 = Rarely; 2 = Maybe/Sometimes; 3 = Probably/Usually; 4 = Yes/Always

	Potential Target Behaviors			
Prioritization Criteria	**(1)** _______	**(2)** _______	**(3)** _______	**(4)** _______
Does this behavior pose danger to the individual or to others?	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
How many opportunities will the individual have to use this new skill in the natural environment? or How often does problem behavior occur?	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
How long-standing is the problem or skill deficit?	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
Will changing this behavior produce a higher rate of reinforcement for the individual?	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
What is the relative importance of this target behavior to future skill development and independent functioning?	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
Will changing this behavior reduce negative or unwanted attention from others?	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
Will changing this behavior produce reinforcement for significant others?	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
How likely it success in changing this behavior?	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
How much will it cost to change this behavior?	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
Totals	_______	_______	_______	_______

and selecting learning goals for students with disabilities. Each behavior is given a number representing the behavior's value on each of the prioritizing variables (e.g., 0 to 4, with 0 representing no value or contribution and 4 representing maximum value or benefit).

Professionals planning behavior change programs for senior citizens would probably insist that target behaviors with immediate benefits receive high priority. Educators serving secondary students with disabilities would likely advocate for behaviors that focus on skill development and independent functioning.

At times, the behavior analyst, the client, and/or significant others have conflicting goals. Parents may want their adolescent daughter in the house by 10:30 PM on weekends, but the daughter may want to stay out until midnight. The school may want a behavior analyst to develop a program to increase students' adherence to an unpopular dress code. The behavior analyst may believe that elements of the dress codes are negotiable, but not to the point that the administration wants the code enforced. Who decides what is best for whom?

One way to minimize conflicts is to obtain client, parent, and staff/administration participation in the goal determination process. For example, the active participation of parents and students, when possible, in the selection of short- and long-term goals is desirable. Such participation by all of the significant parties can avoid and resolve goal conflicts, not to mention the invaluable information the participants can provide relative to other aspects of program planning (e.g., identification of likely reinforcers). Reviewing the results of assessment efforts and allowing each participant to provide input on the relative merits of each proposed goal or target behavior can often produce consensus on the best direction. Program planners should not commit *a priori* that whatever behavior is ranked first will necessarily be considered the highest priority target behavior. However, if the important people involved in a person's life go through a ranking process such as the one shown in Figure 3.9, they are likely to identify areas of agreement and disagreement, which leads to further discussions of target behavior selection and concentration on the critical concerns of those involved.

DEFINING TARGET BEHAVIORS

Before a behavior undergoes analysis, it must be defined in a clear, objective, and concise manner. In constructing target behavior definitions, applied behavior analysts must consider the functional and topographical implications of their definitions.

Role and Importance of Target Behavior Definitions in Applied Behavior Analysis

Applied behavior analysis derives its validity from its systematic approach to seeking and organizing knowledge about human behavior. Validity of scientific knowledge in its most basic form implies replication. When predicted behavioral effects can be reproduced, principles of behavior are confirmed and methods of practice developed. If applied behavior analysts employ definitions of behavior not available to other scientists, replication is less likely. Without replication, the usefulness or meaningfulness of data cannot be determined beyond the specific participants themselves, thereby limiting the orderly development of the discipline as a useful technology (Baer, Wolf, & Risley, 1968). Without explicit, well-written definitions of target behaviors, researchers would be unable to accurately and reliably measure the same response classes within and across studies; or to aggregate, compare, and interpret their data.[7]

Explicit, well-written definitions of target behavior are also necessary for the practitioner, who may not be so concerned with replication by others or development of the field. Most behavior analysis programs are not conducted primarily for the advancement of the field; they are implemented by educators, clinicians, and other human services professionals to improve the lives of their clients. However, implicit in the application of behavior analysis is an accurate, ongoing evaluation of the target behavior, for which an explicit definition of behavior is a must.

A practitioner concerned only with evaluating his efforts to provide optimum service to his clients might ask, "As long as I know what I mean by [name of target behavior], why must I write down a specific definition?" First, a good behavioral definition is operational. It provides the opportunity to obtain complete information about the behavior's occurrence and nonoccurrence, and it enables the practitioner to apply procedures in a consistently accurate and timely fashion. Second, a good definition increases the likelihood of an accurate and believable evaluation of the program's effectiveness. Not only does an evaluation need to be accurate to guide ongoing program decisions, but also the data must be believable to those with a vested interest in the program's effectiveness. Thus, even though the practitioner may not be interested in demonstrating an analysis to the field at large, he must always be concerned with demonstrating effectiveness (i.e., accountability) to clients, parents, and administrators.

Two Types of Target Behavior Definitions

Target behaviors can be defined functionally or topographically.

Function-based Definitions

A **function-based definition** designates responses as members of the targeted response class solely by their common effect on the environment. For example, Irvin, Thompson, Turner, and Williams (1998) defined hand mouthing as any behavior that resulted in "contact of the fingers, hand, or wrist with the mouth, lips, or tongue" (p. 377). Figure 3.10 shows several examples of function-based definitions.

Applied behavior analysts should use function-based definitions of target behaviors whenever possible for the following reasons:

- A function-based definition encompasses all relevant forms of the response class. However, target behavior definitions based on a list of specific topographies might omit some relevant members of the response class and/or include irrelevant response topographies. For example, defining children's offers to play with peers in terms of specific things the children say and do might omit

Figure 3.10 Function-based definitions of various target behaviors.

Creativity In Children's Blockbuilding

The child behaviors of blockbuilding were defined according to their products, *block forms*. The researchers created a list of 20 arbitrary, but frequently seen forms, including:

Arch—any placement of a block atop two lower blocks not in contiguity.

Ramp—a block leaned against another, or a triangular block placed contiguous to another, to simulate a ramp.

Story—two or more blocks placed one atop another, the upper block(s) resting solely upon the lower.

Tower—any story of two or more blocks in which the lowest block is at least twice as tall as it is wide. (Goetz & Baer, 1973, pp. 210–211).

Exercise by Obese Boys

Riding a stationary bicycle—each wheel revolution constituted a response, which was automatically recorded by magnetic counters (DeLuca & Holborn, 1992, p. 672).

Compliance at Stop Signs by Motorists

Coming to a complete stop—observers scored a vehicle as coming to a complete stop if the tires stopped rolling prior to the vehicle entering the intersection (Van Houten & Retting, 2001, p. 187).

Recycling by Office Employees

Recycling office paper—number of pounds and ounces of recyclable office paper found in recycling and trash containers. All types of paper accepted as recyclable were identified as well as examples of nonrecyclable paper (Brothers, Krantz, & McClannahan, 1994, p. 155).

Safety Skills to Prevent Gun Play by Children

Touching the firearm—the child making contact with the firearm with any part of his or her body or with any object (e.g., a toy) resulting in the displacement of the firearm.

Leaving the area—the child removing himself or herself from the room in which the firearm was located within 10 seconds of seeing the firearm (Himle, Miltenberger, Flessner, & Gatheridge, 2004, p. 3).

responses to which peers respond with reciprocal play and/or include behaviors that peers reject.

- The outcome, or function, of behavior is most important. This holds true even for target behaviors for which form or aesthetics is central to their being valued as socially significant. For example, the flowing strokes of the calligrapher's pen and the gymnast's elegant movements during a floor routine are important (i.e., have been selected) because of their effects or function on others (e.g., praise from the calligraphy teacher, high scores from gymnastics judges).
- Functional definitions are often simpler and more concise than topography-based definitions, which leads to easier and more accurate and reliable measurement and sets the occasion for the consistent application of intervention. For example, in their study on skill execution by college football players, Ward and Carnes (2002) recorded a correct tackle according to the clear and simple definition, "if the offensive ball carrier was stopped" (p. 3).

Function-based definitions can also be used in some situations in which the behavior analyst does not have direct and reliable access to the natural outcome of the target behavior, or cannot use the natural outcome of the target behavior for ethical or safety reasons. In such cases, a *function-based definition by proxy* can be considered. For example, the natural outcome of elopement (i.e., running or walking away from a caregiver without consent) is a lost child. By defining elopement as "any movement away from the therapist more than 1.5 m without permission" (p. 240), Tarbox, Wallace, and Williams (2003) were able to measure and treat this socially significant target behavior in a safe and meaningful manner.

Topography-based Definitions

A **topography-based definition** identifies instances of the target behavior by the shape or form of the behavior. Topography-based definitions should be used when the behavior analyst (a) does not have direct, reliable, or easy access to the functional outcome of the target behavior and/or (b) cannot rely on the function of the behavior because each instance of the target behavior does not produce the relevant outcome in the natural environment or the outcome might be produced by other events. For example, Silvestri (2004) defined and measured two classes of positive teacher statements according to the words that made up the statements, not according to whether the comments produced specific outcomes (see Figure 3.11).

Figure 3.11 Topography-based definitions for two types of teacher statements.

Generic Positive Statements

Generic positive statements were defined as audible statements by the teacher that referred to one or more student's behavior or work products as desirable or commendable (e.g., "I'm proud of you!", "Great job, everyone."). Statements made to other adults in the room were recorded if they were loud enough to be heard by the students and made direct reference to student behavior or work products (e.g., "Aren't you impressed at how quietly my students are working today?"). A series of positive comments that specified neither student names nor behaviors with less than 2 seconds between comments was recorded as one statement. For example, if the teacher said, "Good, good, good. I'm so impressed" when reviewing three or four students' work, it was recorded as one statement.

Teacher utterances not recorded as generic positive statements included (a) statements that referred to specific behavior or student names. (b) neutral statements indicating only that an academic response was correct (e.g., "Okay", "Correct"), (c) positive statements not related to student behavior (e.g., saying "Thanks for dropping off my attendance forms" to a colleague), and (d) incomprehensible or inaudible statements.

Behavior-Specific Positive Statements

Behavior-specific positive statements made explicit reference to an observable behavior (e.g., "Thank you for putting your pencil away"). Specific positive statements could refer to general classroom behavior (e.g., "You did a great job walking back to your seat quietly") or academic performance (e.g., "That was a super smart answer!"). To be recorded as separate responses, specific positive statements were separated from one another by 2 seconds or by differentiation of the behavior praised. In other words, if a teacher named a desirable behavior and then listed multiple students who were demonstrating the behavior, this would be recorded as one statement (e.g., "Marissa, Tony, and Mark, you did a great job of returning your materials when you were finished with them"). However, a teacher's positive comment noting several different behaviors would be recorded as multiple statements regardless of the Interval between the end of one comment and the start of the next. For example, "Jade, you did a great job *cleaning up* so quickly; Charles, thanks for *putting the workbooks away*; and class, I appreciate that you *lined up quietly*" would be recorded as three positive statements.

Adapted from *The Effects of Self-Scoring on Teachers' Positive Statements during Classroom Instruction* (pp. 48–49) by S. M. Silvestri. Unpublished doctoral dissertation. Columbus, OH: The Ohio State University. Used by permission.

Topography-based definitions can also be used for target behaviors for which the relevant outcome is sometimes produced in the natural environment by undesirable variations of the response class. For example, because a duffer's very poor swing of a golf club sometimes produces a good outcome (i.e., the ball lands on the green), it is better to define a correct swing by the position and movement of the golf club and the golfer's feet, hips, head, and hands.

A topography-based definition should encompass all response forms that would typically produce the relevant outcome in the natural environment. Although topography provides an important element for *defining* target behaviors, the applied behavior analyst must be especially careful not to *select* target behaviors solely on the basis of topography (see Box 3.2).

Writing Target Behavior Definitions

A good definition of a target behavior provides an accurate, complete, and concise description of the behavior to be changed (and therefore measured). It also states what is not included in the behavioral definition. Asking aloud to be excused from the dinner table is an observable and measurable behavior that can be counted. By comparison, "exercising good manners" is not a description of any particular behavior; it merely implies a general response class of polite and socially acceptable behaviors. Hawkins and Dobes (1977) described three characteristics of a good target behavior definition that are as valid and useful today as they were when published more than 40 years ago:

1. The definition should be *objective*, referring only to observable characteristics of the behavior (and environment, if needed) or translating any inferential terms (such as "expressing hostile feelings," "intended to help," or "showing interest in") into more objective ones.
2. The definition should be *clear* in that it should be readable and unambiguous so that experienced observers could read it and readily paraphrase it accurately.
3. The definition should be *complete*, delineating the "boundaries" of what is to be included as an instance of the response and what is to be excluded, thereby directing the observers in all situations that are likely to occur and leaving little to their judgment. (p. 169)

Stated succinctly, first, a good definition must be objective, ensuring that specific instances of the defined target behavior are observed and recorded reliably. An objective definition increases the likelihood of an accurate and believable evaluation of program effectiveness. Second, a clear definition is technological, meaning that it enables others to use and replicate it (Baer et al., 1968). A clear definition therefore becomes operational for present and future purposes. Finally, a complete definition discriminates between what is and what is not an instance of the target behavior. A complete definition allows others to record an occurrence of a target behavior, but

BOX 3.2

How Serious Are These Behavior Problems?

Suppose you are a behavior analyst in a position to design and help implement an intervention to change the following four behaviors:

1. A child repeatedly raises her arm, extending and retracting her fingers toward her palm in a gripping/releasing type of motion.
2. An adult with developmental disabilities pushes his hand hard against his eye, making a fist and rubbing his eye rapidly with his knuckles.
3. Several times each day a high school student rhythmically drums her fingers up and down, sometimes in bursts of 10 to 15 minutes in duration.
4. A person repeatedly grabs at and squeezes another person's arms and legs so hard that the other person winces and says "Ouch!"

How much of a problem does the behavior pose for the person or for others who share his or her current and future environments? Would you rate each behavior as a mild, moderate, or serious problem? How important do you think it would be to target each of these behaviors for reduction or elimination from the repertoires of the four individuals?

Appropriate answers to these questions cannot be found in topographical descriptions alone. The meaning and relative importance of any operant behavior can be determined only in the context of the environmental antecedents and consequences that define the behavior. Here is what each of the four people in the previous examples were actually doing:

1. An infant learning to wave "bye-bye."
2. A man with allergies rubbing his eye to relieve the itching.
3. A student typing unpredictable text to increase her keyboarding fluency and endurance.
4. A massage therapist giving a relaxing, deep-muscle massage to a grateful and happy customer.

Applied behavior analysts must remember that the meaning of any behavior is determined by its function, not its form. Behaviors should not be targeted for change on the basis of topography alone.

Note: Examples 1 and 2 are adapted from Meyer and Evans. 1989. p. 53.

not record instances of nonoccurrence, in a standard fashion. A complete definition is a precise and concise description of the behavior of interest. Note how the target behavior definitions in Figures 3.10 and 3.11 meet the standard for being objective, clear, and complete.

Morris (1985) suggested testing the definition of a target behavior by asking three questions:

1. Can you count the number of times that the behavior occurs in, for example, a 15-minute period, a 1-hour period, or 1 day? Or, can you count the number of minutes that it takes for the child to perform the behavior? That is, can you tell someone that the behavior occurred "x" number of times or "x" number of minutes today? (Your answer should be "yes.")
2. Will a stranger know exactly what to look for when you tell him or her the target behavior you are planning to modify? That is, can you actually see the child performing the behavior when it occurs? (Your answer should be "yes.")
3. Can you break down the target behavior into smaller behavioral components, each of which is more specific and observable than the original target behavior? (Your answer should be "no.")

In responding to the suggestion that perhaps a sourcebook of standard target behavior definitions be developed because it would increase the likelihood of exact replications among applied researchers and would save the considerable time spent in developing and testing situation-specific definitions, Baer (1985) offered the following perspectives. Applied behavior analysis programs are implemented because someone (e.g., teacher, parent, individual himself) has "complained" that a behavior needs to be changed. A behavioral definition has validity in applied behavior analysis only if it enables observers to capture every aspect of the behavior that the complainer is concerned with and none other. Thus, to be valid from an applied perspective, definitions of target behaviors should be situation specific. Attempts to standardize behavior definitions assume an unlikely similarity across all situations.

SETTING CRITERIA FOR BEHAVIOR CHANGE

Target behaviors are selected for study in applied behavior analysis because of their importance to the people involved. Applied behavior analysts attempt to increase, maintain, and generalize adaptive, desirable behaviors and decrease the occurrence of maladaptive, undesirable behaviors. Behavior analysis efforts that not only target important behaviors but also change those behaviors in a positive and meaningful way are said to have **social validity**.[8] But how much does a target behavior need to change before it makes a meaningful difference in the person's life?

Van Houten (1979) made a case for specifying the desired outcome criteria before efforts to modify the target behavior begin.

> This step [specifying outcome criteria] becomes as important as the previous step [selecting socially important target behaviors] if one considers that for most behaviors there exists a range of responding within which performance is

> most adaptive. When the limits of this range are unknown for a particular behavior, it is possible to terminate treatment when performance is above or below these limits. Hence, the behavior would not be occurring within its optimal range. . . .
>
> In order to know when to initiate and terminate a treatment, practitioners require socially validated standards for which they can aim. (pp. 582, 583)

Van Houten (1979) suggested two basic approaches to determining socially valid goals: (a) assess the performance of people judged to be highly competent and (b) experimentally manipulate different levels of performance to determine empirically which produces optimal results.

Regardless of the method used, specifying treatment goals before intervention begins provides a guideline for continuing or terminating a treatment. Further, setting objective, predetermined goals helps to eliminate disagreements or biases among those involved in evaluating a program's effectiveness.

SUMMARY

Role of Assessment in Applied Behavior Analysis

1. Behavioral assessment involves using indirect, direct, and empirical methods to identify, define, and determine the function of target behaviors.
2. Behavioral assessment consists of five phases or functions: (a) screening, (b) defining and quantifying problems or goals, (c) pinpointing the target behavior(s) to be treated, (d) monitoring progress, and (e) following up.
3. Before conducting a behavioral assessment, the behavior analyst must determine whether she has the authority and permission, resources, and skills to assess and change the behavior. Past and current records related to medical, educational, and historical events should be examined and analyzed as part of a complete behavioral assessment.

Assessment Methods Used by Behavior Analysts

4. Subsumed under the three major methods for assessment information—indirect, direct, and empirical—are (a) interviews, checklists/rating scales, (b) tests and direct observations, and (c) functional behavior analysis and reinforcer/punishment preference assessment methods, respectively.
5. The client interview is used to determine the client's description of problem behaviors or achievement goals. *What, when*, and *where* questions are emphasized, focusing on the actual behavior of the client and the responses of significant others to that behavior.
6. Questionnaires and needs assessment surveys are sometimes completed by the client to supplement the information gathered in the interview.
7. Clients are sometimes asked to self-monitor certain situations or behaviors. Self-collected data may be useful in selecting and defining target behaviors.
8. Significant others can also be interviewed to gather assessment information and, in some cases, to find out whether they will be willing and able to assist in an intervention.
9. Direct observation with a behavior checklist that contains specific descriptions of various skills can indicate possible target behaviors.
10. Anecdotal observation, also called ABC recording, yields a descriptive, temporally sequenced account of all behaviors of interest and the antecedent conditions and consequences for those behaviors as those events occur in the client's natural environment.
11. Ecological assessment entails gathering a large amount of information about the person and the environments in which that person lives and works (e.g., physiological conditions, physical aspects of the environment, interactions with others, past reinforcement history). A complete ecological assessment is neither necessary nor warranted for most applied behavior analysis programs.
12. Reactivity refers to the effects that an assessment procedure has on the behavior being assessed. Behavior analysts should use assessment methods that are as unobtrusive as possible, repeat observations until apparent reactive effects subside, and take possible reactive effects into account when interpreting the results of observations.
13. Choose assessment methods that produce reliable and valid results, conduct assessments according to professional standards, and apply conservative analyses when interpreting results.

Assessing the Social Significance of Potential Target Behaviors

14. Target behaviors in applied behavior analysis must be socially significant behaviors that will increase a person's habilitation (adjustment, competence, and quality of life).
15. The relative social significance and habilitative value of a potential target behavior can be clarified by viewing it in light of the following considerations:
    - Will the behavior be reinforced in the person's daily life? The relevance of behavior rule requires that a target behavior produce reinforcement for the person in the postintervention environment.

- Is the behavior a necessary prerequisite for a useful skill?
- Will the behavior increase the person's access to environments in which other important behaviors can be learned or used?
- Will the behavior predispose others to interact with the person in a more appropriate and supportive manner?
- Is the behavior a cusp or pivotal behavior? Behavioral cusps have sudden and dramatic consequences that extend well beyond the idiosyncratic change itself because they expose the person to new environments, reinforcers, contingencies, responses, and stimulus controls. Learning a pivotal behavior produces corresponding modifications or covariations in other untrained behaviors.
- Is the behavior age appropriate?
- Whenever a behavior is targeted for reduction or elimination, a desirable, adaptive behavior must be selected to replace it.
- Does the behavior represent the actual problem or achievement goal, or is it only indirectly related?
- A person's verbal behavior should not be confused with the actual behavior of interest. However, in some situations the client's verbal behavior should be selected as the target behavior because it is the behavior of interest.
- If a person's goal is not a specific behavior, a target behavior(s) must be selected that will produce the desired results or state.

Prioritizing Target Behaviors

16. Assessment often reveals more than one possible behavior or skill area for targeting. Prioritization can be accomplished by rating potential target behavior against key questions related to their relative danger, frequency, long-standing existence, potential for reinforcement, relevance for future skill development and independent functioning, reduced negative attention from others, likelihood of success, and cost.
17. Participation by the person whose behavior is to be changed, parents and/or other important family members, staff, and administration in identifying and prioritizing target behaviors can help reduce goal conflicts.

Defining Target Behaviors

18. Explicit, well-written target behavior definitions are necessary for researchers to accurately and reliably measure the same response classes within and across studies or to aggregate, compare, and interpret their data.
19. Good target behaviors definitions are necessary for practitioners to collect accurate and believable data to guide ongoing program decisions, apply procedures consistently, and provide accountability to clients, parents, and administrators.
20. Function-based definitions designate responses as members of the targeted response class solely by their common effect on the environment.
21. Topography-based definitions define instances of the targeted response class behavior by the shape or form of the behavior.
22. A good definition must be objective, clear, and complete, and must discriminate between what is and what is not an instance of the target behavior.
23. A target behavior definition is valid if it enables observers to capture every aspect of the behavior that the "complainer" is concerned with and none other.

Setting Criteria for Behavior Change

24. A behavior change has social validity if it changes some aspect of the person's life in an important way.
25. Outcome criteria specifying the extent of behavior change desired or needed should be determined before efforts to modify the target behavior begin.
26. Two approaches to determining socially validated performance criteria are (a) assessing the performance of people judged to be highly competent and (b) experimentally manipulating different levels of performance to determine which produces optimal results.

KEY TERMS

ABC recording
anecdotal observation
behavior checklist
behavioral assessment
behavioral cusp
ecological assessment
function-based definition
habilitation
normalization
pivotal behavior
reactivity
relevance of behavior rule
social validity
target behavior
topography-based definition

MULTIPLE-CHOICE QUESTIONS

1. All of the following are assessment methods used in applied behavior analysis except:
   a. Direct observation
   b. Interview
   c. Medical evaluations
   d. Checklists

   Hint: (See "Role of Assessment in Applied Behavior Analysis")

2. Behavior assessment seeks to determine the ______ of a behavior.
   a. Function
   b. Replacement
   c. Importance
   d. Structure

   Hint: (See "Role of Assessment in Applied Behavior Analysis")

3. All of the following are competencies of a behavior analyst except:
   a. Knowledge of socially important behavior
   b. Technical skills
   c. Ability to conduct statistical analyses of data
   d. Ability to match assessment data with intervention strategies

   Hint: (See "Phases of Behavioral Assessment")

4. Potential target behaviors should not be selected if the primary reason for selection is the:
   a. Benefit of others
   b. Safety of the person
   c. Safety of family members
   d. Potential to increase independence

   Hint: (See "Assessing the Social Significance of Potential Target Behaviors")

5. Interview questions should avoid "why" questions because these tend to encourage _______ explanations of behavior.
   a. Defensive
   b. Mentalistic
   c. Lengthy
   d. Direct

   Hint: (See "Interviewing the Client")

6. A descriptive and temporally sequenced account of behavior in the natural environment is called a(n):
   a. Anecdotal observation
   b. Behavioral assessment
   c. Ecological assessment
   d. Real time observation

   Hint: (See "Direct Observations")

7. The ________ of behavior rule states that a target behavior should only be selected when it can be determined that the behavior will produce natural reinforcement.
   a. Functionality
   b. Validity
   c. Relevance
   d. Importance

   Hint: (See "Will This Behavior Likely to Produce Reinforcement in the Client's Natural Environment after Treatment Ends?")

8. The principle of ________ determines the degree to which a person's behavior repertoire maximizes short and long-term reinforcers for that individual and for others, and minimizes short and long-term punishers.
   a. Normalization
   b. Habilitation
   c. Functionality
   d. Justification

   Hint: (See "Assessing the Social Significance of Potential Target Behaviors")

9. Behavior that exposes an individual to new contingencies, reinforcers, and stimulus controls is called:
   a. Pivotal behavior
   b. Access behavior
   c. Behavior cusp
   d. Contingent behavior

   Hint: (See "Is This Behavior a Behavioral Cusp or a Pivotal Behavior?")

10. When a problem behavior has been targeted for reduction or elimination, the behavior analyst must always include a(n) ________ in the intervention plan.
    a. Replacement behavior
    b. Appropriate response
    c. Pivotal behavior
    d. Performance criterion

    Hint: (See "If the Proposed Target Behavior is to be Reduced or Eliminated, What Adaptive Behavior Will Replace It?")

11. Juan is a six-year-old boy with a developmental disability who attends an integrated kindergarten class. Assessments have identified four target behaviors. Which behavior should be the first target for intervention?
    a. Flicking his fingers in front of his eyes
    b. Bolting from the playground
    c. Toilet training
    d. Humming loudly during group activities

    Hint: (See "Prioritizing Target Behaviors")

12. In determining the likelihood of success in changing a behavior, all of the following should be considered except:

a. Research on changing this behavior

b. Experience of the behavior analyst

c. Social validity of the behavior

d. Available resources

Hint: (See "Prioritizing Target Behaviors")

13. Explicit behavior definitions are important in research of applied behavior analysis for all of the following except:

a. Replication by other scientists

b. Accurate and reliable measurement of behavior

c. Comparison of data across studies

d. Agreement between assessment and intervention data

Hint: (See "Defining Target Behaviors")

14. A(n) ________ definition designates responses in terms of their effect on the environment.

a. Observable

b. Function-based

c. Topography-based

d. Ecological

Hint: (See "Defining Target Behaviors")

15. A behaviors has ________ if it changes some aspect of the person's life in an important way.

a. Social validity

b. Functional application

c. Observable benefit

d. Normalized outcomes

Hint: (See "Setting Criteria for Behavior Change")

16. Outcome criteria should be established before intervention commences for all of the following reasons except:

a. To establish the target performance level

b. To know when to terminate intervention

c. To ensure accurate data collection

d. To ensure agreement on outcomes among stakeholders

Hint: (See "Setting Criteria for Behavior Change")

ESSAY-TYPE QUESTIONS

1. How does behavioral assessment differ from educational or psychological assessment?

Hint: (See "Role of Assessment in Applied Behavior Analysis")

2. Describe one of the competencies of an applied behavior analyst.

Hint: (See "Phases of Behavioral Assessment")

3. Describe one of the four major assessment methods used in behavior analysis.

Hint: (See "Assessment Methods Used by Behavior Analysts")

4. Identify one reason social validity has become a major focus of applied behavior analysis.

Hint: (See "Assessing the Social Significance of Potential Target Behavior")

5. Describe the relevance of behavior rule and briefly explain how it can be used to identify socially significant behavior.

Hint: (See "Habilitation Defined")

6. What is the first question to ask when prioritizing potential target behaviors and why should it be asked first?

Hint: (See "Prioritizing Target Behaviors")

7. Explain one research benefit of explicit, well-written definitions in applied behavior analysis.

Hint: (See "Defining Target Behaviors")

8. What are two features of a function-based behavior definition?

Hint: (See "Two Types of Target Behavior Definitions")

9. Describe one benefit to setting the criteria for behavior change prior to initiating intervention.

Hint: (See "Setting Criteria for Behavior Change")

NOTES

1. Problem behaviors can be assessed empirically using a three-step process called *functional behavior assessment* that is designed to identify and systematically manipulate antecedents and/or consequences that may control the occurrence of the problem behavior(s). Chapter 27 describes this process in detail.
2. Chapter 31, "Ethical and Professional Responsibilities of Applied Behavior Analysts," examines this important issue in detail.
3. Behavior analysts use other forms of assessment to guide the design of effective interventions. Those assessment methods are described elsewhere in this text: preference assessment and reinforcer assessment (Chapter 11), punisher assessment (Chapter 14), and functional behavior assessment (Chapter 27).
4. ABC recording is discussed further in Chapter 27, "Functional Behavior Assessment."
5. Reactive effects of assessment are not necessarily negative. Self-monitoring has become as much a treatment procedure as it is an assessment procedure; see Chapter 29.
6. A target behavior's indirect benefit as a necessary prerequisite for another important behavior should not be confused with indirect teaching. Indirect teaching involves selecting a target behavior different from the true purpose of the behavior because of a belief that they are related (e.g., having a student with poor reading skills practice shape discrimination or balance beam walking). The importance of directness in target behavior selection is discussed later in this section.
7. Procedures for measuring behavior accurately and reliably are discussed in Chapter 4.
8. A third component of social validity concerns the social acceptability of the procedures used to change the behavior. Chapter 10 describes procedures for assessing the social validity of behavior change.

CHAPTER 4

Measuring Behavior

LEARNING OBJECTIVES

- Describe the functions of measurement in applied behavior analysis.
- Identify the measurable dimensions of behavior.
- Describe the different procedures for measuring behavior.
- State the advantages and disadvantages of using continuous measurement procedures and sampling procedures.
- Explain the procedure for measuring behavior by permanent products.
- Explain computer-assisted measurement of behavior.
- Select the appropriate measurement procedure given the dimensions of the behavior and the logistics of observing and recording.

When I awoke this morning, I looked at the clock to see how many hours I slept. (Do I stay in bed or get up? I got up.) I showered, shaved, brushed my teeth, and then stepped onto the bathroom scale to check my weight. (Did I weigh the same, less, or more than yesterday? I gained two pounds.) I started thinking about how much we humans use measurement in our everyday activities, and for the most part we are completely oblivious of the behavioral measurement.

For breakfast I made coffee (4 scoops of coffee, with 4½ cups of water) and placed an English muffin in the toaster (set the toaster to the correct temperature for toasting the muffin). I had breakfast and then I was off to the community rec center for morning exercise.

I got in the car, and the speedometer told me how many miles per hour I was driving. (I stayed within the speed limit.) The gas gauge told me how much gas was in the tank. (Enough to drive to the rec center, return home, and then some.)

At the rec center, I observed people walking or running on the track. One walker recorded the number of laps on a hand-held tally counter, a runner glanced at the clock after each lap, and several walkers and runners wore digital step counters on their wrists. In other places in the rec center, people counted and recorded the number of repetitions they completed with free weights and strength-building machines.

And on and on it goes throughout the day. Measurement makes up a large portion of our everyday lives. It greatly improves the quality of our interaction with our environment. Just as measurement is so important in our environment, it is the foundation of all activity in applied behavior analysis.

Measurement provides the basis for all scientific discoveries and for the development and successful application of technologies derived from those discoveries. Direct and frequent measurement constitutes the foundation for applied behavior analysis. Applied behavior analysts use measurement to detect and compare the effects of various environmental arrangements on the acquisition, maintenance, and generalization of socially significant behaviors.

What is it about behavior that applied behavior analysts can and should measure? How should those measures be obtained? And, what should we do with these measures once we have obtained them? This chapter identifies the dimensions by which behavior can be measured and describes the measurement methods behavior analysts commonly use. First, we examine the definition and functions of measurement in applied behavior analysis.

DEFINITION AND FUNCTIONS OF MEASUREMENT IN APPLIED BEHAVIOR ANALYSIS

Measurement is the process of applying quantitative labels to describe and differentiate objects and natural events. "[It] involves attaching a number representing the observed extent of a dimensional quantity to an appropriate unit. The number

and the unit together constitute the measure of the object or event" (Johnston & Pennypacker, 1993a, p. 95). Measurement in applied behavior analysis entails three steps: (a) identifying the behavior to be measured, (b) defining the behavior in observable terms, and (c) selecting an appropriate observation and data-recording method (Gast, 2014). Chapter 3 covered selecting and defining target behavior; this chapter details measurement methods.

Bloom, Fischer, and Orme (2003)—who described measurement as the act or process of applying quantitative or qualitative labels to events, phenomena, or observed properties using a standard set of consensus-based rules by which to apply labels to those occurrences—pointed out that the concept of measurement includes the characteristics of what is being measured, the quality and appropriateness of the measurement tools, the technical skill of the measurer, and how the measures obtained are used. In the end, measurement gives researchers, practitioners, and consumers a common means for describing and comparing behavior with a set of labels that convey a common meaning.

Researchers Need Measurement

> *Dr. Watson: "What do you imagine that means?"*
> *Sherlock Holmes: "I have no data yet. It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts."*
> —from A Scandal in Bohemia by Arthur Conan Doyle

Measurement is how scientists operationalize empiricism. Objective measurement enables (indeed, it requires) scientists to describe the phenomena they observe in precise, consistent, and publicly verifiable ways. Without measurement, all three levels of scientific knowledge—description, prediction, and control—would be relegated to guesswork subject to the "individual prejudices, tastes, and private opinions of the scientist" (Zuriff, 1985, p. 9). We would live in a world in which the alchemist's suppositions about a life-prolonging elixir would prevail over the chemist's formulated compounds derived from experimentation.

Behavior analysts measure behavior to answer questions about the existence and nature of functional relations between socially significant behavior and environmental variables. Measurement enables comparisons of a person's behavior within and between different environmental conditions, thereby affording the possibility of drawing empirically based conclusions about the effects of those conditions on behavior. For example, Dunlap and colleagues (1994) measured students' task engagement and disruptive behaviors during choice and no-choice conditions. Measurement revealed the level of both target behaviors during each condition, whether and how much the behaviors changed when choice was introduced or withdrawn, and how variable or stable the behaviors were during each condition.

The researcher's ability to achieve a scientific understanding of behavior change depends on her ability to measure it. Measurement makes possible the detection and verification of virtually everything that has been discovered about the effects of the environment on behavior. The empirical databases of the basic and applied branches of behavior analysis consist of organized collections of behavioral measurements. Virtually every graph in the *Journal of Applied Behavior Analysis* and *Journal of the Experimental Analysis of Behavior* displays an ongoing record or summary of behavioral measurement. In short, measurement provides the very basis for learning and talking about behavior in scientifically meaningful ways.[1]

Practitioners Need Measurement

Behavioral practitioners are dedicated to improving their clients' lives by changing socially significant behaviors. Practitioners measure behavior initially to determine the current level of a target behavior and whether that level merits intervention. If intervention is warranted, the practitioner measures the extent to which his efforts are successful. Practitioners measure behavior to find out whether and when it has changed; the extent and duration of behavior changes; the variability or stability of behavior before, during, and after treatment; and whether important behavior changes have occurred in other settings or situations and spread to other behaviors.

Frequent measures of behavior during treatment (*formative assessment*) enable dynamic, data-based decision making concerning the continuation, modification, or termination of treatment. Practitioners also compare measurements of the target behavior before and after treatment (sometimes including pre- and post-treatment measures obtained in nontreatment settings or situations) to evaluate the overall effects of behavior change programs (*summative evaluation*).

The practitioner who does not obtain and attend to frequent measures of the target behavior is vulnerable to committing mistakes: (a) continuing an ineffective treatment when no real behavior change has occurred or (b) discontinuing an effective treatment because subjective judgment detects no improvement (e.g., without measurement, a teacher would be unlikely to know that a student's oral reading rate has increased from 70 to 80 words per minute). Direct and frequent measurement enables practitioners to detect their successes and, equally important, their failures so they can make changes to transform failure into success (Bushell & Baer, 1994; Greenwood & Maheady, 1997; Gast, 2014).

> Our technology of behavior change is also a technology of behavior measurement and of experimental design; it developed as that package, and as long as it stays in that package, it is a self-evaluating enterprise. Its successes are successes of known magnitude; its failures are almost immediately detected as failures; and whatever its outcomes, they are attributable to known inputs and procedures rather than to chance events or coincidences. (D. M. Baer, October 21, 1982)

In addition to enabling ongoing program monitoring and data-based decision making, frequent measurement

provides other important benefits to practitioners and the clients they serve:

- *Measurement helps practitioners optimize their effectiveness.* To be optimally effective, a practitioner must maximize behavior change efficiency in terms of time and resources. Only by maintaining close, continual contact with relevant outcome data can a practitioner hope to achieve optimal effectiveness and efficiency (Bushell & Baer, 1994). Commenting on the critical role of direct and frequent measurement in maximizing the effectiveness of classroom practice, Sidman (2000) noted that teachers "must remain attuned to the pupil's messages and be ready to try and to evaluate modifications [in instructional methods]. Teaching, then, is not just a matter of changing the behavior of pupils; it is an interactive social process" (p. 23, words in brackets added). Direct and frequent measurement is the process by which practitioners hear their clients' messages.
- *Measurement enables practitioners to verify the legitimacy of treatments touted as "evidence based."* Practitioners are increasingly expected, and in some fields mandated by law, to use evidence-based interventions. An evidence-based practice is a treatment or intervention that has been shown to be effective by a substantial body of high-quality, peer-reviewed scientific research. When implementing any treatment, regardless of the type or amount of research evidence to support it, practitioners can and should verify its effectiveness with the students or clients they serve by direct and frequent measurement.
- *Measurement helps practitioners identify and terminate treatments based on pseudoscience, fad, fashion, or ideology.* Many controversial treatments and proposed cures for people with developmental disabilities and autism (e.g., facilitated communication, holding therapy, megadoses of vitamins, strange diets, weighted vests, dolphin-assisted therapy) have been promoted in the absence of sound scientific evidence of effectiveness (Foxx & Mulick, 2016). The use of so-called breakthrough therapies has led to disappointment and loss of precious instructional or therapeutic time and, in some cases, to disastrous consequences (Maurice, 1993; Todd, 2012). Even though well-controlled studies have shown many of these methods to be ineffective, and even though these programs are not justified because they lacked sound scientific evidence of effects, risks, and benefits, parents and practitioners are still bombarded with sincere and well-meaning testimonials. Measurement is the practitioner's best ally in the quest to find and verify effective treatments, and to root out those whose strongest support consists of testimonials and slick Internet ads. Practitioners should maintain a healthy skepticism regarding claims for effectiveness. See Box 31.5 for further discussion on this topic.

  Using Plato's Allegory of the Cave as a metaphor for teachers and other practitioners who use untested and pseudo-instructional ideas, Heron, Tincani, Peterson, and Miller (2005) argued that practitioners would be better served by casting aside pseudo-educational theories and philosophies and adopting a scientific approach. Practitioners who directly and frequently measure the effects of their intervention and treatment programs have empirical support to defend against political or social pressures to adopt unproven treatments. In a real sense they arm themselves with what Carl Sagan (1996) called a "baloney detection kit."
- *Measurement helps practitioners be accountable to clients, consumers, employers, and society.* Practitioners should use direct and frequent measurement of behavior to answer parent and caregiver questions about the development of their children or clients.
- *Measurement helps practitioners achieve ethical standards.* Ethical codes of conduct for behavior analytic practitioners require direct and frequent measurement of client behavior (Behavior Analyst Certification Board, 2018; Chapter 31 this text). Determining whether a client's right to effective treatment is being honored requires measurement of the behavior(s) for which treatment was sought or intended. Failing to measure the nature and extent of relevant behavior changes of clients borders on malpractice. Kauffman (2005) offered this perspective on the relationship between measurement and ethical practice in education:

  > [T]he teacher who cannot or will not pinpoint and measure the relevant behaviors of the students he or she is teaching is probably not going to be very effective. . . . Not to define precisely and to measure these behavioral excesses and deficiencies, then, is a fundamental error; it is akin to the malpractice of a nurse who decides not to measure vital signs (heart rate, respiration rate, temperature, and blood pressure), perhaps arguing that he or she is too busy, that subjective estimates of vital signs are quite adequate, that vital signs are only superficial estimates of the patient's health, or that vital signs do not signify the nature of the underlying pathology. The teaching profession is dedicated to the task of changing behavior—changing behavior demonstrably for the better. What can one say, then, of educational practice that does not include precise definition and reliable measurement of the behavioral change induced by the teacher's methodology? It is indefensible. (p. 439)

MEASURABLE DIMENSIONS OF BEHAVIOR

If a friend asked you to measure a coffee table, you would probably ask why he wants the table measured. In other words, what does he want measurement to tell him about the table? Does he need to know its height, width, and depth? Does he want to know how much the table weighs? Perhaps he is interested in the color of the table? Each of these reasons for measuring the table requires measuring a different dimensional quantity of the table (e.g., length, mass, light reflection).

Behavior, like coffee tables and all entities in the physical world, also has features that can be measured. Because behavior

occurs within and across time, it has three fundamental and measurable *dimensional quantities*. Johnston and Pennypacker (2009) described these dimensional quantities as:

- **Repeatability** (also called *countability*): Instances of a behavior can occur repeatedly through time (i.e., behavior can be counted).
- **Temporal extent**: Every instance of behavior occurs during some amount of time (i.e., the duration of behavior can be measured).
- **Temporal locus**: Every instance of behavior occurs at a certain point in time with respect to other events (i.e., when behavior occurs can be measured).

Figure 4.1 shows a schematic representation of repeatability, temporal extent, and temporal locus. Alone and in combination, these dimensional quantities provide the basic and derivative measures used by applied behavior analysts. In the following pages these and two other measurable dimensions of behavior—its form and strength—will be discussed.

Measures Based on Repeatability

Count

Count is a simple tally of the number of occurrences of a behavior. Although how often a behavior occurs is often of primary interest, measures of count alone may not provide enough information to allow behavior analysts to make useful program decisions or analyses. For example, data showing that Katie wrote correct answers to 5, 10, and 15 long division problems over three consecutive math class periods suggests improving performance. However, if the three measures of count were obtained in observation periods of 5 minutes, 20 minutes, and 60 minutes, respectively, a much different interpretation of Katie's performance is suggested. Therefore, the observation period, or counting time, should always be noted when reporting measures of count.

Rate

Combining count and observation time yields rate, one of the most widely used measures in behavior analysis. Skinner (1953) considered rate of response the basic measurement for behavioral research. He invented the cumulative recorder, a device that automatically produced a graphic record of operant response rates by his experimental subjects (see Figure 1.1).

Rate is defined as the number of responses per unit of time.[2] A rate measure is a ratio consisting of the dimensional quantity of count (number of responses) and time (observation period in which the count was obtained). Converting count to rate of response makes measurement more meaningful. For example, knowing that Yumi read 95 words correctly and 4 words incorrectly in 1 minute, that Lee wrote 250 words in 10 minutes, and that Joan's self-injurious behavior occurred 17 times in 1 hour provides important information and context. Expressing the previously reported counts of Katie's performance in math class as rate reveals that she correctly answered long division problems at rates of 1.0, 0.5, and 0.25 per minute over three consecutive class periods.

Rate is often reported as a count per 30 seconds, count per minute, count per hour, count per day—or occasionally, count per week, count per month, or count per year. As long as the unit of time is standard within or across experiments, rate measures can be compared. For example, a student who, over four daily class activities of different durations, had 12 talk-outs in 20 minutes, 8 talk-outs in 12 minutes, 9 talk-outs in 15 minutes, and 12 talk-outs in 18 minutes had response rates of 0.60, 0.67, 0.60, and 0.67 per minute.

The six rules and guidelines that follow will help researchers and practitioners obtain, describe, and interpret count and rate data most appropriately.

Reference the Counting Time. Behavior analysts use two methods to calculate rate of response: (a) total observation time and (b) total interresponse time (IRT) (Johnston & Pennypacker, 2009). Applied behavior analysts most often use the total observation time to calculate rate. To calculate total IRT, the analysts first measure the IRT for each response (i.e., the amount of time that elapses between two consecutive instances of a behavior) and then calculate the sum of the IRTs. (See the section on IRT later in this chapter.)

When using rate of response, behavior analysts must include the duration of the observation time. Comparing rate measures without reference to the counting time can lead to faulty interpretations of data. Consider, for example, that Sally and Lillian each ran at a rate of 7 minutes per mile. We cannot compare their performances without reference to the distances they ran. Running 1 mile at a rate of 7 minutes per mile is a different class of behavior than running a marathon (26.2 miles) at a rate of 7 minutes per mile.

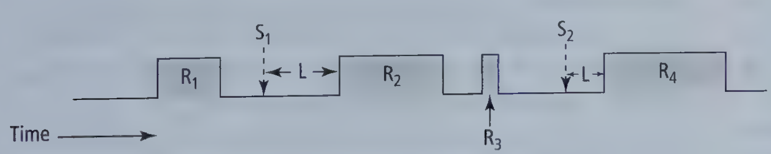


Figure 4.1 Schematic representation of the dimensional quantities of repeatability, temporal extent, and temporal locus. *Repeatability* is shown by a count of four instances of a given response class (R_1, R_2, R_3, and R_4) within the observation period. The *temporal extent* (i.e., duration) of each response is represented by the raised and shaded portions of the time line. One aspect of the *temporal locus* (response latency) of two responses (R_2 and R_4) is represented by the elapsed time (←L→) between the onset of two antecedent stimulus events (S_1 and S_2) and the initiation of the responses that follow.

The counting time used for each session needs to accompany each rate measure when the counting time changes from session to session. For instance, rather than having a set counting time to answer arithmetic facts (e.g., a 1-minute timing), the teacher records the total time required for the student to complete an assigned set of arithmetic problems during each class session. In this situation, the teacher could report the student's correct and incorrect answers per minute for each session, and also report the counting times for each session because they changed from session to session.

Calculate Correct and Incorrect Rates of Response When Assessing Skill Development. When a participant has an opportunity to make correct and incorrect responses, a rate of response for each behavior should be reported. Calculating rates of correct and incorrect responses is crucial for evaluating skill development because knowing only the correct rate does not allow the behavior analyst or participant to assess an improving performance. The rate of correct responses alone could show an improving performance, but if the rate of incorrect responding is also increasing, the improvement may be illusionary. Correct and incorrect rate measures together provide important information to help the teacher evaluate how well the student is progressing. Ideally, the correct response rate accelerates toward a performance criterion, while the incorrect response rate decelerates to a performance criterion. Also, reporting rate correct and rate incorrect enables an assessment of proportional accuracy while maintaining the dimensional quantities of the measurement (e.g., 20 correct and 5 incorrect responses per minute = 80% accuracy, or a multiple of ×4 proportional accuracy).

Correct and incorrect response rates provide essential data for assessing fluent performance (i.e., proficiency) (Kubina, 2005). The assessment of fluency requires measuring the number of correct and incorrect responses per unit of time (i.e., proportional accuracy). Fluency cannot be evaluated using correct rate only because fluent performance must be accurate also.

Consider the Varied Complexity of Responses. Rate of responding is a sensitive and appropriate measure of skill acquisition and the development of fluent performances only when the level of difficulty and complexity from one response to the next remains constant within and across observations. The rates of response previously discussed have been with whole units in which the response requirements are essentially the same from one response to the next. Many important behaviors, however, are composites of two or more component behaviors, and different situations call for varied sequences or combinations of the component behaviors.

One method for measuring rate of responding that takes varied complexity into account for multiple-component behaviors is to count the operations required to achieve a correct response. For example, in measuring students' math calculation performance, instead of counting answers to a two-digit plus three-digit addition problem with regrouping as correct or incorrect, a behavior analyst might consider the number of steps that were completed in correct sequence within each problem. Helwig (1973) used the number of operations needed to produce the answers to mathematics problems to calculate response rates. Each session, students were given 20 multiplication and division problems selected at random from a set of 120 problems. The teacher recorded the duration of time for each session. All the problems were of two types: $a \times b = c$ and $a \div b = c$. For each problem, the students were asked to find one of the factors: the product, the dividend, the divisor, or the quotient. Depending on the problem, finding the missing factor required from one to five operations. For example, writing the answer 275 in response to the problem $55 \times 5 = ?$ would be scored as four correct responses because finding the missing factor requires four operations:

1. Multiply the ones: $5 \times 5 = 25$.
2. Record the 5 ones and carry the 2 tens.
3. Multiply the tens: $5 \times 5(0) = 25(0)$.
4. Add the 2 tens carried and record the sum (27).

If multiple ways to find the answer were possible, the mean number of operations was counted. For example, the answer to the problem $4 \times ? = 164$ can be obtained by two operations with multiplication and by four operations with division. The mean number of operations is three. Helwig counted the number of operations completed correctly and incorrectly per set of 20 problems and reported correct and incorrect rates of response.

Use Rate to Measure Free Operants. Rate of response is a useful measure for all behaviors characterized as free operants. The term **free operant** refers to behaviors that have discrete onsets and offsets, do not depend on discriminative stimuli, involve minimal displacement of the organism in time and space (upon completion of one response, the organism is immediately ready to emit another response in the class), and can be emitted over a wide range of response rates. Skinner (1966) used rate of response of free operants as the primary dependent variable when developing the experimental analysis of behavior. The bar press by rats and key peck by pigeons are typical free operant responses used in nonhuman animal laboratory studies. Many socially significant behaviors meet the definition of free operants: words read during a 1-minute counting period, head slaps per minute, letter strokes written in 3 minutes.

Rate of response is a preferred measurement for free operants because it is sensitive to changes in behavior values (e.g., oral reading may occur at rates ranging from 0 to 250 or more correct words per minute) and because it offers clarity and precision by defining a count per unit of time.

Do Not Use Rate to Measure Behaviors That Occur Within Discrete Trials. Rate of response is an inappropriate measure for behaviors that can occur only within limited or restricted situations. For example, response rates of behaviors that occur within **discrete trials** are controlled by a given opportunity to emit the response. Typical discrete trials used in nonhuman animal laboratory studies include moving from one end of a maze

or shuttle box to another. Applied examples of discrete trials include responding to teacher-presented flash cards; answering a question prompted by the teacher; and, when presented with a sample color, pointing to a color from an array of three colors that matches the sample color. In each of these examples, rate of response is controlled by the presentation of a discriminative stimulus. Because behaviors that occur within discrete trials are opportunity bound, measures such as percentage of response opportunities in which a response was emitted or trials-to-criterion should be used.

Do Not Use Rate to Measure Continuous Behaviors That Occur for Extended Periods. Rate is also a poor measure for continuous behaviors that occur for extended periods, such as participating in a playground game or being on task during a classroom activity. Such behaviors are best measured by whether they are "on" or "off" at any given time, yielding data on duration or estimates of duration obtained by interval recording.

Celeration

Just like a car that goes faster when the driver presses the gas pedal and slows down when the driver releases the pedal or steps on the brake, rates of response accelerate and decelerate. **Celeration**, the root word of acceleration and deceleration, is a measure of how rates of response change over time. Rate accelerates when a participant responds faster over successive counting periods and decelerates when responding slows over successive observations. Applied behavior analysts should use the terms *acceleration* or *deceleration* when describing increasing or decreasing rates of response.

Celeration is a measure of count per unit time/per unit of time; or expressed another way, rate/per unit of time (Graf & Lindsley, 2002; Kubina, 2005). Celeration—changes in rate—is a direct measure of dynamic patterns of behavior change, such as transitions from one steady state of responding to another and the acquisition of fluent levels of performance (Cooper, 2005). The Standard Celeration Chart provides a standard format for displaying measures of celeration. There are four Standard Celeration Charts, showing rate as count (a) per day, (b) per week, (c) per month, and (d) per year. These four charts provide different levels of magnification for viewing and interpreting celeration. Methods for plotting and interpreting celeration data on the Standard Celeration Chart are described in Chapter 6.

Measures Based on Temporal Extent

Duration

Duration, the amount of time from the onset to the end point of a response, is the basic measure of temporal extent. Applied behavior analysts measure duration in standard units of time (e.g., Enrique worked cooperatively with his peer tutor for 6 minutes and 24 seconds today).

Duration is important when measuring the amount of time a person engages in a target behavior. Applied behavior analysts measure the duration of target behaviors that a person has been engaging in for too long or for too short of a time period, such as a child with developmental disabilities who has tantrums for more than an hour at a time, or a student who sticks with an academic task for no more than 30 seconds.

Duration is also an appropriate measure for behaviors that occur at very high rates (e.g., rocking, rapid jerks of the head, hands, legs) or task-oriented continuous behaviors that occur for an extended time (e.g., cooperative play, on-task behavior, off-task behavior).

Behavioral researchers and practitioners commonly measure one or both of two kinds of duration measures: *total duration* per session or observation period and *duration per occurrence*.

Total Duration per Session. Total duration is a measure of the cumulative time a person engages in the target behavior. Applied behavior analysts use two procedures to measure and report total duration. One method involves recording the cumulative amount of time a target behavior occurs during an observation period. An observer measuring the total time a child engages in solitary play during free-play periods activates a stopwatch when the child begins to play alone. The moment solitary play ceases, the observer stops the stopwatch but does not reset it. When the child drifts into solitary play again, the observer restarts the stopwatch. The observer continues to start and stop the stopwatch, matching the child's starting and ending bouts of solitary play. If the duration of the observation periods remains constant (e.g., 10 minutes), total duration per session data can be reported in standard units of time (e.g., 6 min and 30 sec of free play). If the duration of observation periods varies, total duration per session data must be converted to a percentage of total time observed (e.g., 6 min and 30 sec of solitary play in a 10-min session = 65%).

Zhou, Iwata, Goff, and Shore (2001) used total duration measurement to assess leisure-item preferences of people with profound developmental disabilities. They used a stopwatch to record physical engagement with an item (i.e., contact between both hands and the item) during 2-minute trials. They reported total contact in seconds by summing the duration values across three 2-minute trials of each assessment. McCord, Iwata, Galensky, Ellingson, and Thomson (2001) measured the total duration in seconds that two adults with severe or profound intellectual disabilities engaged in problem behavior (see Figure 6.6).

The other measure of total duration recording is the amount of time a person spends completing a specific task, without specifying a minimum or maximum observation period—for example, the number of minutes a person takes getting washed and dressed for work, after arising in the morning, or the time needed to complete a flat tire repair by a high school student enrolled in a work-study program.

Duration per Occurrence. Duration per occurrence is a measure of the duration of time that each instance of the target behavior occurs. For example, Greene, Bailey, and Barber (1981) used a sound-recording device to automatically record (a) the number of times noisy disruptions by children on a school bus exceeded a specified sound threshold and (b) the duration in seconds that each disruption remained above that

threshold. The researchers used the mean duration per occurrence of noisy disruptions as one measure for evaluating the intervention's effects.

Selecting and Combining Measures of Count and Duration. Measurements of count, total duration, and duration per occurrence provide different views of behavior. Count and duration measure different dimensional quantities of behavior, and these differences provide the basis for selecting which dimension(s) to measure. Count measures repeatability, whereas duration recording measures temporal extent. For instance, a teacher concerned about a student who is out of her seat "too much" could tally each time the student leaves her seat. The behavior is discrete and is unlikely to occur at such a high rate that counting the number of occurrences would be difficult. Because any instance of out-of-seat behavior has the potential to occur for an extended time and the total time the student is out of her seat is a socially significant aspect of the behavior, the teacher could also use total duration recording.

Using count to measure out-of-seat behavior provides the number of times the student left her seat. A measure of total duration will indicate the amount and proportion of time that the student was out of her seat during the observation period. Because of the relevance of temporal extent in this case, duration would be a better measure than count. The teacher might observe that the student left her seat once in a 30-minute observation period. One occurrence in 30 minutes might not be viewed as a problem. However, if the student left her seat at minute 1 of the observation period and did not return for the remainder of the period, a very different view of the behavior is obtained.

In this situation, duration per occurrence would make a better measurement selection than count or total duration recording because duration per occurrence measures repeatability (i.e., the rate of response) *and* the temporal extent of the behavior. A duration-per-occurrence measure will give the behavior analyst information on the number of times the student was out of her seat and the duration of each occurrence. Duration per occurrence is often preferable to total duration because rate is sensitive to the number of instances and duration of the target behavior. Further, if a total duration measure is useful for other purposes, the individual durations of each of the counted and timed occurrences can be summed. However, if behavior endurance (e.g., academic responding, motor movements) is the major consideration, then total duration recording may be sufficient (e.g., 5 minutes of oral reading, 10 minutes of free writing).

Measures Based on Temporal Locus

Temporal locus refers to (a) when a behavior occurs with respect to other events of interest (e.g., the latency between the onset of an antecedent event and the occurrence of the behavior) and (b) the amount of time that elapses between two consecutive instances of a response class. These two points of reference provide the context for measuring response latency and interresponse time (IRT), the two measures of temporal locus most frequently reported in the behavior analysis literature.

Latency

Latency is a measure of the elapsed time between the onset of a stimulus and a subsequent response.[3] Latency is an appropriate measure when the applied behavior analyst is interested in how much time occurs between the opportunities to emit a behavior and when the subject initiates the target behavior. For example, Edwards, La Londe, Cox, Weetjens, and Poling (2016) researched the effects of schedules of reinforcement (see Chapter 13) on rats' search for living people trapped under rubble. "Rats were exposed to 10 trials each day. For each trial, the release point for rats and the location of the two human targets, who assumed seated positions within the rubble, were selected from random lists" (p. 200). The researchers recorded (a) the latency from the rat's search-area release point to finding and placing both front paws on a prepositioned human target under simulated rubble and (b) after an initiation buzzer sounded, a second latency measure from the target back to the release point. This significant research emphasized the importance of quickly locating and providing treatment to living people who have been trapped under rubble.

Interest can also focus on latencies that are too short. A student may answer incorrectly because she does not wait for the teacher to complete the questions. An adolescent who, at the slightest provocation from a peer, immediately retaliates has no time to consider alternative behaviors that could defuse the situation and lead to improved interactions.

Behavior analysts typically report response latency data as the mean, median, and range of individual latency measures per observation period. For example, Lerman, Kelley, Vorndran, Kuhn, and LaRue (2002) used a latency measure to assess the effects of different reinforcer magnitudes (i.e., 20 seconds, 60 seconds, or 300 seconds of access to a reinforcer) on post-reinforcement pause—the absence of responding for a period following reinforcement. The researchers measured the number of seconds from the end of each reinforcer-access interval to the first instance of the target behavior (a communication response). They then calculated and graphed the mean, median, and range of response latencies measured during each session (see Lerman et al., 2002, p. 41).

Interresponse Time

Interresponse time (IRT) is a measure of the amount of time that elapses between two consecutive instances of a behavior. Like response latency, IRT is a measure of temporal locus because it identifies *when* a specific instance of behavior occurs with respect to another event (i.e., the previous response). Figure 4.2 shows a schematic representation of interresponse time.

Although a direct measure of temporal locus, IRT is functionally related to rate of response. Shorter IRTs coexist with higher rates of response, and longer IRTs occur within lower response rates. Applied behavior analysts measure IRT when the time between instances of a response class is important. IRT provides a basic measure for implementing and evaluating interventions using differential reinforcement of low and high rates (DRL & DRH), a procedure for using reinforcement to reduce (i.e., DRL) or increase (i.e., DRH) the rate of

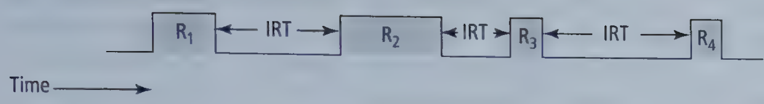


Figure 4.2 Schematic representation of three interresponse times (IRT). IRT, the elapsed time between the termination of one response and the initiation of the next response, is a commonly used measure of temporal locus.

responding (see Chapter 25). Like latency data, IRT measures are typically reported and displayed graphically by mean (or median) and range per observation period.

IRT measures appear frequently in basic behavior analysis experiments examining the effects of schedules of reinforcement (Tanno, Silberberg, & Sakagami, 2009) and variables related to conditioned reinforcement (Bejarano & Hackenberg, 2007), delayed reinforcement (Lattal & Ziegler, 1982), and punishment (Galbicka & Platt, 1984). In one of the few applied studies featuring IRT measures, Lennox, Miltenberger, and Donnelly (1987) reduced rapid eating by three adults with profound intellectual disabilities. The intervention combined differential reinforcement of low rate responding (DRL; see Chapter 25) (an IRT of 15 sec was required before an eating response was permitted), response interruption (blockage of eating responses before 15 seconds had elapsed since the previous response), and prompting participants to engage in an incompatible behavior (putting their forks down and hands in their laps) during the 15-second interval.

Derivative Measures

Behavior analysts frequently use *percentage* and *trials-to-criterion*, two forms of data derived from direct measures of dimensional quantities of behavior.

Percentage

Percentage is a ratio (i.e., a proportion) formed by combining the same dimensional quantities, such as count (i.e., a tally divided by a tally) or time (i.e., a duration divided by a duration; a latency divided by a latency). Percentage expresses the proportional quantity of some event in terms of the number of times the event occurred per 100 opportunities that the event could have occurred. For example, if a student correctly answered 39 of 50 exam items, an accuracy percentage would be calculated by dividing the number of correct answers by the total number of items and multiplying that product by 100 ($39 \div 50 = 0.78 \times 100 = 78\%$).

Percentage is frequently used in applied behavior analysis to report participants' response accuracy. For example, Ward and Carnes (2002) used a percentage-of-correct-performances measure in their study evaluating the effects of goal setting and public posting on skill execution of three defensive skills by linebackers on a college football team. The researchers recorded counts of correct and incorrect reads, drops, and tackles by each player and calculated accuracy percentages based on the number of opportunities for each type of play. (Data from this study are shown in Figure 9.3.) Petursdottir and Aguilar (2016) reported percentage correct data on three kindergarten boys' acquisition of receptive identification tasks as a function of the order of stimulus presentation in matching-to-sample trials.

Applied behavior analysts often use percentage to report the proportion of observation intervals in which the target behavior occurred. These measures are typically reported as a percentage of intervals within a session (e.g., see Figures 6.4 and 12.7). Percentage can also be calculated for an entire observation session. In a study analyzing the differential effects of reinforcer quality, immediacy, and response effort on the impulsive behavior of students with attention-deficit/hyperactivity disorder, Neef, Bicard, and Endo (2001) reported the percentage of time each student allocated to two sets of concurrently available math problems (e.g., time allocated to math problems yielding high-quality delayed reinforcers ÷ total time possible × 100 = %).

Percentages are used widely in education, psychology, and the popular media, and most people understand proportional relationships expressed as percentages. However, percentages are often used improperly. Accordingly, we offer several notes of caution on the use and interpretation of percentages.

Percentages most accurately reflect the level of and changes in behavior when calculated with a divisor (or denominator) of 100 or more. However, most percentages used by behavior analysts are calculated with divisors much smaller than 100. Percentage measures based on small divisors are unduly affected by small changes in behavior. For example, a change in count of just 1 response per 10 opportunities changes the percentage by 10%. Guilford (1965) cautioned that it is unwise to compute percentages with divisors smaller than 20. For research purposes, we recommend that whenever possible, applied behavior analysts design measurement systems in which resultant percentages are based on no fewer than 30 response opportunities or observation intervals.

Changes in percentage may erroneously suggest improving performance. For example, an accuracy percentage could increase even though the occurrence of incorrect responses remains the same or worsens. Consider a student whose accuracy in answering math problems on Monday is 50% (5 of 10 problems answered correctly) and on Tuesday is 60% (12 of 20 problems answered correctly). Even with the improved proportional accuracy, the number of errors increased (from 5 on Monday to 8 on Tuesday).

Although no other measure communicates proportional relationships better than percentage, its use as a behavioral quantity is limited because a percentage has no dimensional quantities.[4] For example, percentage cannot be used to assess the development of proficient or fluent behavior because an assessment of proficiency must reference count and time. Percentage can, however, show the proportional accuracy of a targeted behavior during the development of proficiency.

Another limitation of percentage as a measure of behavior change is that it imposes lower and upper limits on the data. For example, using percent correct to assess a student's

reading performance establishes an artificial ceiling on the measurement of performance. A learner who correctly reads 100% of the words she is presented cannot improve according to the measure used.

Different percentages can be reported from the same data set, with each percentage suggesting significantly different interpretations. For example, consider a student who scores 4 correct (20%) on a 20-item pretest and 16 correct (80%) when the same 20 items are given as a posttest. The most straightforward description of the student's improvement from pretest to posttest (60%) compares the two measures using the original basis or divisor (20 items). Because the student scored 12 more items correct on the posttest than he did on the pretest, his performance on the posttest could be reported as an increase (gain score) over his pretest performance of 60%. And given that the student's posttest score represented a fourfold improvement in correct responses, some might report the posttest score as a 300% improvement of the pretest—a completely different interpretation from an improvement of 40%.

Although percentages greater than 100% are sometimes reported, strictly speaking, doing so is incorrect. Although a behavior change greater than 100% may seem impressive, it is a mathematical impossibility. A percentage is a *proportional* measure of a total set, where *x* (the proportion) of *y* (the total) is expressed as 1 part in 100. A proportion of something cannot exceed the total of that something or be less than zero (i.e., there is no such thing as a negative percentage). Every coach's favorite athlete who "always gives 110%" simply does not exist.[5]

Trials-to-Criterion

Trials-to-criterion is a measure of the number of response opportunities needed to achieve a predetermined level of performance. What constitutes a trial depends on the nature of the target behavior, the relevant setting variables and antecedent stimuli, and the desired performance level. For a skill such as shoe tying, each opportunity to tie a shoe could be considered a trial, and trials-to-criterion data are reported as the number of trials required for the learner to tie a shoe correctly without prompts or assistance. For behaviors involving problem solving or discriminations that must be applied across a large number of examples to be useful, a trial might consist of a block, or series of response opportunities, in which each response opportunity involves the presentation of a different exemplar of the problem or discrimination. For example, a trial for discriminating between the short and long vowel sounds of the letter *o* could be a block of 10 consecutive opportunities to respond, in which each response opportunity is the presentation of a word containing the letter *o*, with short-vowel and long-vowel *o* words (e.g., *hot, boat*) presented in random order. Trials-to-criterion data could be reported as the number of blocks of 10 trials required for the learner to correctly pronounce the *o* sound in all 10 words. Count would be the basic measure from which the trials-to-criterion data would be derived.

Behavior analysts can also use other basic measures as trials-to-criterion data (i.e., rate, duration, latency). For example, a trials-to-criterion measure for solving two-digit minus two-digit subtraction problems requiring borrowing could be the number of practice sheets of 20 randomly generated and sequenced problems a learner completes before she is able to solve all 20 problems on a single sheet in 3 minutes or less.

Trials-to-criterion data are often calculated and reported as an ex post facto measure of one important aspect of the "cost" of a treatment or instructional method. For example, Trask-Tyler, Grossi, and Heward (1994) reported the number of instructional trials needed by each of three students with visual impairments and developmental disabilities to prepare three recipes without assistance on two consecutive times over two sessions. Each recipe entailed from 10 to 21 task-analyzed steps.

Trials-to-criterion data are used frequently to compare the relative efficiency of two or more treatments or instructional methods. For example, by comparing the number of practice trials needed for a student to master weekly sets of spelling words practiced in two different ways, a teacher could determine whether the student learns spelling words more efficiently with one method than with another. Sometimes trials-to-criterion data are supplemented by information on the number of minutes of instruction needed to reach predetermined performance criteria (e.g., Holcombe, Wolery, Werts, & Hrenkevich 1993; Repp, Karsh, Johnson, & Van Laarhoven, 1994).

Trials-to-criterion measures can also be collected and analyzed as a dependent variable throughout a study. For example, R. Baer (1987) recorded and graphed trials-to-criterion on a paired-associates memory task as a dependent variable in a study assessing the effects of caffeine on the behavior of preschool children.

Trials-to-criterion data can also be useful for assessing a learner's increasing competence in acquiring a related class of concepts. For instance, teaching a concept such as the color red to a child could consist of presenting "red" and "not red" items to the child and differentially reinforcing correct responses. Trials-to-criterion data could consist of the number of "red" and "not red" exemplars required before the child achieves a specified level of performance with the discrimination. The same instructional and data collection procedures could then be used in teaching other colors to the child. Data showing that the child achieves mastery of each newly introduced color in fewer instructional trials than required to learn previous colors might be evidence of the child's increasing agility in learning color concepts.

Definitional Measures

In addition to the basic and derived dimensions already discussed, behavior can also be defined and measured by its form and intensity. Neither form (i.e., topography) nor intensity (i.e., magnitude) of responding is a fundamental dimensional quantity of behavior, but each is an important parameter for defining and verifying the occurrence of many response classes. Behavior analysts measure the topography or magnitude of a response to determine whether the response represents an occurrence of the target behavior. Occurrences of the target behavior verified on the basis of topography or magnitude are then measured by one or more aspects of count, temporal extent, or temporal locus. In other words, measuring topography or magnitude is sometimes necessary to determine whether instances of the targeted

response class have occurred, but the subsequent quantification of those responses is recorded, reported, and analyzed in terms of the fundamental and derivative measures of count, rate, duration, latency, IRT, percentage, and trials-to-criterion.

Topography

Topography, the physical form or shape of a behavior, is a measurable and malleable dimension of behavior. Topography is a measurable dimension of behavior because responses of varying form can be detected from one another. That topography is a malleable aspect of behavior is evidenced by the fact that responses of varying form are shaped and selected by their consequences.

A group of responses with widely different topographies may serve the same function (i.e., form a response class). For example, each of the different ways of writing the word *topography* shown in Figure 4.3 would produce the same effect on most readers. Membership in some response classes, however, is limited to responses within a narrow range of topographies. Although each of the response topographies in Figure 4.3 would meet the functional requirements of most written communications, none would meet the standards required of an advanced calligraphy student.

Topography is of obvious and primary importance in activities in which form, style, or artfulness of behavior is valued in its own right (e.g., painting, sculpting, dancing, gymnastics). Measuring and providing differential consequences for responses of varied topographies is also important when the functional outcomes of the behavior correlate highly with specific topographies. A student who sits with good posture and looks at the teacher is more likely to receive positive attention and opportunities to participate academically than is a student who slouches, with her head on the desk (Schwarz & Hawkins, 1970). Basketball players who execute foul shots with a certain form make a higher percentage of shots than when they shoot idiosyncratically (Kladopoulos & McComas, 2001; see Figure 6.3).

Trap, Milner-Davis, Joseph, and Cooper (1978) measured the topography of cursive handwriting by first-grade students. Plastic transparent overlays were used to detect deviations from model letters in lower- and uppercase letters written by the children (see Figure 4.4). The researchers counted the number of correct letter strokes—those that met all of the specified topographical criteria (e.g., all letter strokes contained within the 2-mm parameters of the overlay, connected, complete, sufficient length)—and used the percentage correct of all strokes

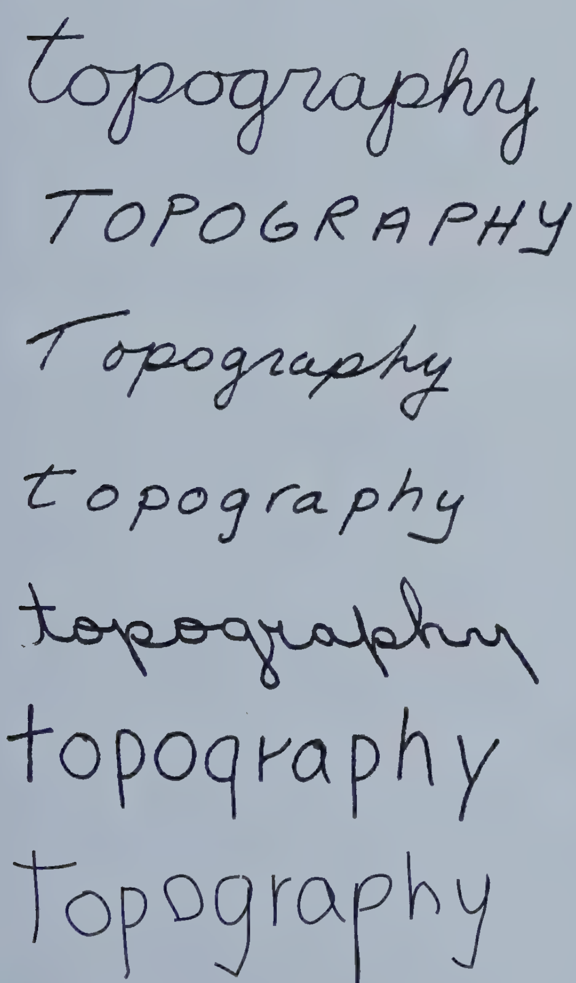


Figure 4.3 Topography, the physical form or shape of behavior, is a measurable dimension of behavior.

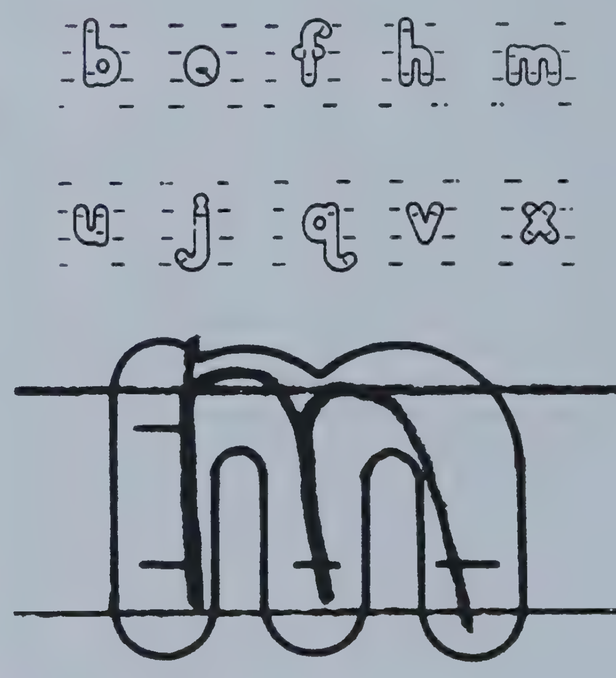


Figure 4.4 Examples of outlines on a transparent overlay used to measure inside and outside boundaries of manuscript letters and an illustration of using the transparent overlay to measure the letter *m*. Because the vertical stroke of the letter *m* extends beyond the outline boundary, it did not meet the topographical criteria for a correct response.

From "The Measurement of Manuscript Letter Strokes" by J. J. Helwig, J. C. Johns, J. E. Norman, J. O. Cooper, 1976. *Journal of Applied Behavior Analysis, 9*, p. 231. Copyright 1976 by the Society for the Experimental Analysis of Behavior, Inc. Used by permission.

written by each student to assess the effects of visual and verbal feedback and a certificate of achievement on the children's acquisition of cursive handwriting skills.

Magnitude

Magnitude refers to the force or intensity of a response. The desired outcomes of some behaviors are contingent on responding at or above (or below) a certain intensity or force. A screwdriver must be turned with sufficient force to insert or remove screws; a pencil must be applied to paper with enough force to leave legible marks. In contrast, applying too much torque to a misaligned screw or bolt is likely to strip the threads, and pushing too hard with a pencil will break its point.

Several studies have measured the magnitude of speech or other vocalizations that were considered too loud or too soft (Koegel & Frea, 1993). Schwarz and Hawkins (1970) measured the voice loudness of Karen, a sixth-grade girl who spoke so softly in class that her voice was usually inaudible to others. Karen's voice was recorded on videotape during two class periods each day. (The videotape was also used to obtain data on two other behaviors: face touching and the amount of time Karen sat in a slouched position). The researchers then played the videotape into an audiotape recorder with a loudness indicator and counted the number of times the needle went above a specified level on the loudness meter. Schwarz and Hawkins used the number (proportion) of needle inflections per 100 words spoken by Karen as the primary measure for evaluating the effects of an intervention on increasing her voice volume during class.

Edgerton and Wine (2017) used a digital tablet and app (Voice Meter Pro™) to measure the voice volume of a boy with autism who frequently spoke below conversational volume. The app displayed a thermometer gauge that rose and fell with the surrounding sound level and signaled changes in voice volume with different background colors, an animated figure, and text prompts ("Speak up, I can't hear you!", "That's better", and "That's too loud!"). The investigators calibrated the app to detect three levels of magnitude (too soft, conversational, and too loud) and recorded the number of times the boy's responses to questions were spoken at conversational volume.

Greene, Bailey, and Barber (1981) measured the magnitude of noisy disruptions by middle school students on a school bus with an automated sound-recording device. The recording device could be adjusted so that only sound levels above a predetermined threshold activated it. The device automatically recorded both the number of times outbursts of sound exceeded a specified threshold (93 dB) and the total duration in seconds that the sound remained above that threshold. When the noise level exceeded the specified threshold, a light on a panel that all students could see was activated automatically. When the light was off, students listened to music during the bus ride; when the number of noisy disruptions was below a criterion, they participated in a raffle for prizes. This intervention drastically reduced the outbursts and other problem behaviors as well. Greene and colleagues reported both the number and the mean duration per occurrence of noisy disruptions as measures of evaluating the intervention's effects.[6] Table 4.1 summarizes the measurable dimensions of behavior and considerations for their use.

METHODS FOR MEASURING BEHAVIOR

Methods for measuring behavior used most often by applied behavior analysts involve one or a combination of the following: event recording, timing, and various time sampling methods.

Event Recording

Event recording encompasses a wide variety of procedures for detecting and recording the number of times a behavior of interest occurs. For example, Cuvo, Lerch, Leurquin, Gaffaney, and Poppen (1998) used event recording to measure the effects of work requirements and reinforcement schedules on the choice behavior of adults with intellectual disabilities and preschool children while they were engaged in age-appropriate tasks (e.g., adults sorting silverware, children tossing beanbags or jumping hurdles). The researchers recorded each piece of silverware sorted, each beanbag tossed, and each hurdle jumped.

Event recording is also used to measure discrete trial behaviors, in which the count for each trial or opportunity to respond is either 1 or 0, representing the occurrence or nonoccurrence of the target behavior. Figure 4.5 shows a form used to record the occurrence of imitation responses by a preschooler with disabilities and his typically developing peer partner within a series of instructional trials embedded into ongoing classroom activities (Valk, 2003). For each trial, the observer recorded the occurrence of a correct response, no response, an approximation, or an inappropriate response by the target child and the peer by circling or marking a slash through letters representing each behavior. The form also allowed the observer to record whether the teacher prompted or praised the target child's imitative behavior.

Considerations for Event Recording. Many applied behavior analysts in practice use non-automated event recording. It is easy to do. Most people can tally discrete behaviors accurately, often on the first attempt. If the response rate is not too high, non-automated event recording may not interfere with other activities. For example, a teacher can continue with instruction while tallying occurrences of the target behavior. Event recording provides useful data for most behaviors. However, each instance of the target behavior must have discrete beginning and ending points. Event recording is applicable for target behaviors such as students' oral responses to questions, students' written answers to math problems, and a parent praising a son or daughter's behavior. Behaviors such as humming are hard to measure with event recording because an observer would have difficulty determining when one hum ends and another begins. Event recording is difficult for behaviors defined without specific discrete action or object relations, such as engagement with materials during free-play activity. Because engagement with materials does not present a specific discrete action or object relation, an observer may

TABLE 4.1 Fundamental, derived, and definitional dimensions by which behavior can be measured and described.

Fundamental measures	How calculated	Considerations
Count: The number of responses emitted.	Simple tally of the number of responses observed. • Judah contributed 5 comments to the class discussion.	• Used in calculating rate, celeration, percentage, and trials-to-criterion.
Rate: A ratio of count per observation time; often expressed as count per standard unit of time (e.g., per minute, per hour, per day).	Report number of responses recorded per unit of time in which observations were conducted. • If Judah's comments were counted during a 10-minute class discussion, his rate of responding would be 5 comments per 10 minutes. Often calculated by dividing the number of responses recorded by the number of standard units of time in which observations were conducted. • Judah commented at a rate of 0.5 per minute.	• Minimize faulty interpretations by reporting counting time. • Evaluating skill development and fluency requires measurement of correct and incorrect response rates. • Account for varied complexity and difficulty when calculating response rates. • Rate is the most sensitive measure of changes in repeatability. • Preferred measure for free operants. • Poor measure for behaviors that occur within discrete trials or for behaviors that occur for extended durations. • Most sensitive measure of behavior repeatability.
Celeration: The change (acceleration or deceleration) in rate of responding over time.	Based on count per unit of time (rate)/per unit of time expressed as the factor by which responding is accelerating/decelerating (multiplying or dividing). • A trend line connecting Judah's mean rates of commenting over 4 weeks of 0.1, 0.2, 0.4, and 0.8 comment per minute, respectively, would show an acceleration of ×2 per week.	• Reveals dynamic patterns of behavior change, such as transitions from one steady state to another and acquisition of fluency. • Displayed with a trend line on a Standard Celeration Chart (see Chapter 6). • Minimum of seven measures of rate recommended for calculating.
Duration: The amount of time from the onset to the end point of a response.	*Total duration:* Two methods: (a) Add the individual amounts of time for each response during an observation period; or (b) record the total time the individual is engaged in an activity, or needs to complete a task, without a minimum or maximum observation period. • Judah commented for 1.5 minutes in class today. *Duration per occurrence:* Record duration of time for each instance of the behavior; often reported by mean or median and range of durations per session. • Judah's 5 comments today had a mean duration of 11 seconds, with a range of 3 to 24 seconds.	• Important measure when target behavior is problematic because it occurs for durations that are too long or too short. • Useful measure for behaviors that occur at very high rates and for which accurate event recording is difficult (e.g., finger flicking). • Useful measure for behaviors that do not have discrete beginnings and for which event recording is difficult (e.g., humming). • Useful measure for task-oriented or continuous behaviors (e.g., cooperative play). • Duration per occurrence often preferred over total duration because it includes data on count and total duration.

(continued)

TABLE 4.1 (continued)

Fundamental measures	How calculated	Considerations
		• Use total duration when increasing the endurance of behavior is the goal. • Measuring duration per occurrence entails counting responses, which can be used to calculate rate of responding.
Latency: The elapsed time between the onset of stimulus and the initiation of a subsequent response.	Record time elapsed from the onset of the antecedent stimulus event and the beginning of the response; often reported by mean or median and range of latencies per session. • Judah's comments today had a mean latency of 30 seconds following a peers' comment (range, 5 to 90 seconds).	• Important measure when target behavior is a problem because it is emitted with latencies that are too long or too short. • Decreasing latencies may reveal increasing mastery of some skills.
Interresponse time (IRT): The amount of time that elapses between two consecutive instances of a response class.	Record time elapsed from the end of previous response and the beginning of next response; often reported by mean or median and range of IRTs per session. • Judah's comments today had a median IRT of 2 minutes and a range of 10 seconds to 5 minutes.	• Important measure when the time between responses, or pacing of behavior, is a concern. • Although a measure of temporal locus, IRT is correlated with rate of responding. • An important measure when implementing and evaluating DRL (see Chapter 23).

Derived measures	How calculated	Considerations
Percentage: A proportion, expressed as a number of parts per 100; typically a ratio of the number of responses of a certain type per total number of responses (or opportunities or intervals in which such a response could have occurred).	Divide number of responses meeting specified criteria (e.g., correct responses, responses with minimum IRT, responses of particular topography) by the total number of responses emitted (or response opportunities) and multiply by 100. • 70% of Judah's comments today were relevant to the discussion topic.	• Percentages based on divisors smaller than 20 are unduly influenced by small changes in behavior. Minimum of 30 observation intervals or response opportunities recommended for research. • Change in percentage may erroneously suggest improved performance. • Always report the divisor on which percentage measures are based. • Cannot be used to assess proficiency or fluency. • Imposes upper and lower limits on performance (i.e., cannot exceed 100% or fall below 0%). • Widely different percentages can be reported from the same data set.

TABLE 4.1 (continued)

Fundamental measures	How calculated	Considerations
		• To calculate an overall percentage from percentages based on different denominators (e.g., 90% [9/10], 87.5% [7/8], 33% [1/3], 100% [1/1], divide the total numerators of the component percentages (e.g., 18) by the total denominators (e.g., 18 / 22 = 81.8%). A mean of the percentages themselves yields a different outcome (e.g., 90% + 87.5% + 33% + 100% / 4 = 77.6%).
Trials-to-criterion: Number of responses, instructional trials, or practice opportunities needed to reach a predetermined performance criterion.	Add number of responses or practice trials necessary for learner to achieve specified criterion. • Judah needed 14 blocks of 10 opportunities to comment to achieve the criterion of 8 on-topic comments per 10 opportunities for 2 consecutive sessions.	• Provides an ex post facto description of the "cost" of a treatment or instructional method. • Useful for comparing relative efficiency of different methods of instruction or training. • Useful in assessing changes in the rate at which a learner masters new skills (agility).

Definitional measures	How calculated	Considerations
Topography: The form or shape of behavior.	Used to determine whether responses meet topographical criteria; responses meeting those criteria are measured and reported by one or more fundamental or derivative measures (e.g., percentage of responses meeting topographical criteria). • The plane of the golf club remained within plus or minus 2 degrees from backswing to follow-through on 85% of Amanda's swings.	• Important measure when desired outcomes of behavior are contingent on responses meeting certain topographies. • Important measure for performance areas in which form, style, or artfulness is valued.
Magnitude: The strength, intensity, or force of behavior.	Used to determine whether responses meet magnitude criteria; responses meeting those criteria are measured and reported by one or more fundamental or derivative measures (e.g., count of responses meeting magnitude criteria). • Jill bench pressed 60 pounds 20 times.	• Important measure when desired outcomes of behavior are contingent on responses within a certain range of magnitudes.

Session date: May 21 Session no: 16 Observer: Jennie

Target child: Jordan Peer: Ethan IOA day: YES NO

Target behavior: Place block on structure Condition: 5-sec time delay

Code: C = Correct N = No response A = Approximation I = Inappropriate

Trial	Target child's behavior	Teacher behavior toward target child	Peer's behavior	Teacher praise
1	C N A I	Prompt Praise	C N A I	Praise
2	C N A I	Prompt Praise	C N A I	Praise
3	C N A I	Prompt Praise	C N A I	Praise
4	C N A I	Prompt Praise	C N A I	Praise
5	C N A I	Prompt Praise	C N A I	Praise
6	C N A I	Prompt Praise	C N A I	Praise
7	C N A I	Prompt Praise	C N A I	Praise
8	C N A I	Prompt Praise	C N A I	Praise
9	C N A I	Prompt Praise	C N A I	Praise
10	C N A I	Prompt Praise	C N A I	Praise

No. corrects by target child: 8 No. corrects by peer: 9

****

Target behavior: Place sticker on paper Condition: 5-sec time delay

Trial	Target child's behavior	Teacher behavior toward target child	Peer's behavior	Teacher praise
1	C N A I	Prompt Praise	C N A I	Praise
2	C N A I	Prompt Praise	C N A I	Praise
3	C N A I	Prompt Praise	C N A I	Praise
4	C N A I	Prompt Praise	C N A I	Praise
5	C N A I	Prompt Praise	C N A I	Praise
6	C N A I	Prompt Praise	C N A I	Praise
7	C N A I	Prompt Praise	C N A I	Praise
8	C N A I	Prompt Praise	C N A I	Praise
9	C N A I	Prompt Praise	C N A I	Praise
10	C N A I	Prompt Praise	C N A I	Praise

No. corrects by target child: 4 No. corrects by peer: 8

Figure 4.5 Data collection form for recording the behavior of two children and a teacher during a series of discrete trials.

Adapted from *The Effects of Embedded Instruction Within the Context of a Small Group on the Acquisition of Imitation Skills of Young Children with Disabilities* by J. E. Valk (2003), p. 167. Unpublished doctoral dissertation, The Ohio State University. Used by permission.

have difficulty judging when one engagement starts and ends, and then another engagement begins.

Another consideration with non-automated event recording is that the target behaviors should not occur at such high rates that an observer would have difficulty counting each discrete occurrence accurately. High-rate behaviors that may be difficult to measure with event recording include rapid talking, body rocks, and tapping objects.

Also, event recording does not produce accurate measures for target behaviors that occur for extended time periods, such as staying on task, listening, playing quietly alone, being out of one's seat, or thumb sucking. Task-oriented or continuous behaviors (e.g., being "on task") are examples of target behaviors for which event recording would not be indicated. Classes of continuous behaviors occurring across time are usually not a prime concern of applied behavior analysts. For example, reading per se is of less concern than the number of words read correctly and incorrectly, or the number of reading comprehension questions answered correctly and incorrectly. Similarly, behaviors that demonstrate understanding are more important to measure than "listening behavior," and the number of academic responses a student emits during an independent seatwork period is more important than being on task.

Timing

Behavior analysts use a variety of timing devices and procedures to measure duration, response latency, and IRT.

Timing Duration

Researchers often use semi-automated computer-driven systems for recording durations. Practitioners, however, will most likely use non-automated instruments for recording duration. The most precise non-automated instrument is a digital stopwatch. Practitioners can use wristwatches and wall clocks to measure duration, but the measures obtained will be less precise than those obtained with a stopwatch.

The procedure for recording the total duration of a target behavior per session with a stopwatch is to (a) activate the stopwatch when the behavior starts and (b) stop the watch at the end of the episode. Then, without resetting the stopwatch, the observer starts the stopwatch again at the beginning of the second occurrence of the behavior and stops the watch at the end of the second episode. The observer continues to accumulate the durations of time in this fashion until the end of the observation period, and then transfers the total duration of time showing on the stopwatch to a data sheet.

Timing Latency and IRT

Procedures for measuring latency and IRT are similar to those used to measure duration. Measuring latency requires the precise detection and recording of the time that elapses from the onset of each occurrence of the antecedent stimulus event of interest to the onset of the target behavior. Measuring IRTs requires recording the precise time that elapses from the termination of each occurrence of the target behavior to the onset of the next response.

Time Sampling

Time sampling refers to a variety of methods for observing and recording behavior during intervals or at specific moments in time. The basic procedure involves dividing the observation period into time intervals and then recording the presence or absence of behavior within or at the end of each interval.

Ethologists originally developed time sampling to study the behavior of animals in the field (Altmann, 1974). Because it was not possible or feasible to observe the animals continuously, these scientists arranged systematic schedules of relatively brief but frequent observation intervals. The measures obtained from these "samples" are considered representative of the behavior during the total time period from which they were collected. For example, much of our knowledge about the behavior of chimpanzees and other primates is based on data collected with time sampling observation methods by researchers such as Jane Goodall (1991).

Applied behavior analysts use three forms of time sampling—whole-interval recording, partial-interval recording, and momentary time sampling.[7]

Whole-Interval Recording

Whole-interval recording is often used to measure continuous behaviors (e.g., cooperative play) and behaviors that occur at such high rates that observers have difficulty distinguishing one response from another (e.g., rocking, humming) but can detect whether the behavior is occurring at any given time. The observation period for **whole-interval recording** is divided into a series of brief time intervals (e.g., 5 to 10 seconds). At the end of each interval, the observer records whether the target behavior occurred *throughout* the interval. Data obtained with whole-interval recording usually underestimate the overall percentage of the observation period in which the behavior actually occurred. The longer the observation intervals, the greater the degree to which whole-interval recording may underestimate the actual occurrence of the behavior.

Data collected with whole-interval recording are reported as the percentage of total intervals in which the target behavior was recorded as occurring. Because whole-interval recording represents the proportion of the entire observation period that the person was engaged in the target behavior, whole-interval recording data yield an estimate of total duration. For example, assume a whole-interval observation period consisted of six 10-second intervals (a 1-minute time frame). If the target behavior occurred for four of these 10-second whole intervals and did not occur for the remaining two intervals, the whole-interval recording would yield a total duration estimate of 40 seconds.

Figure 4.6 shows an example of a whole-interval recording form used to measure the on-task behavior of four students during academic seatwork time (Ludwig, 2004). Each minute was divided into four 10-second observation intervals; each observation interval was followed by 5 seconds in which the observer recorded the occurrence or nonoccurrence of target behavior during the preceding 10 seconds. The observer first watched Student 1 continuously for 10 seconds, and then she looked away during the next 5 seconds and recorded whether Student 1 had been on task throughout the previous 10 seconds by circling YES or NO on the recording form. After the 5-second interval for recording Student 1's behavior, the observer looked up and watched Student 2 continuously for 10 seconds, after which she recorded Student 2's behavior on the form. The same procedure for observing and recording was used for Students 3 and 4. In this way, the on-task behavior of each student was observed and recorded for one 10-second interval per minute.

Continuing the sequence of observation and recording intervals over a 30-minute observation period provided thirty 10-second measures (i.e., samples) of each student's on-task behavior. The data in Figure 4.6 show that the observer judged the four students to have been on task during Session 17 for 87%, 93%, 60%, and 73% of the intervals, respectively. Although the data are intended to represent the level of each student's behavior throughout the observation period, it is important to remember that each student was observed for a total of only 5 of the observation period's 30 minutes.

Observers using any form of time sampling should always make a recording response of some sort in every interval. For example, an observer using a form such as the one in Figure 4.6 would record the occurrence or nonoccurrence of the target behavior in each interval by circling YES or NO. Leaving unmarked intervals increases the likelihood of losing one's place on the recording form and marking the result of an observation in the wrong interval space.

All time sampling methods require a timing device to signal the beginning and end of each observation and recording interval. Observers using pencil, paper, clipboard, and timers for interval measurement will often attach a stopwatch to a clipboard. However, observing and recording behavior while having to look simultaneously at a stopwatch is likely to have a negative impact on the accuracy of measurement. As an effective solution to this problem, the observer can listen by earphone to prerecorded audio cues signaling the observation and recording intervals. For example, observers using a whole-interval recording procedure like the one just described could listen to an audio recording with a sequence of prerecorded statements such as the following: "Observe Student 1"; 10 seconds later, "Record Student 1"; 5 seconds later, "Observe Student 2"; 10 seconds later, "Record Student 2"; and so on.

Tactile prompting devices can also be used to signal observation intervals. For example, the *Gentle Reminder* (dan@gentlereminder.com) and *MotivAider* (www.habitchange.com) are small timing instruments that vibrate at the time intervals programmed by the user.

On-Task Recording Form

Date: May 7 Group no: 1 Session no: 17

Observer: Robin IOA session: ___Yes X No

Experimental condition: Baseline On task Productivity

Obs. start time: 9:42 Stop time: 10:12

10-sec intervals	Student 1		Student 2		Student 3		Student 4	
1	YES	NO	YES	NO	YES	NO	YES	NO
2	YES	NO	YES	NO	YES	NO	YES	NO
3	YES	NO	YES	NO	YES	NO	YES	NO
4	YES	NO	YES	NO	YES	NO	YES	NO
5	YES	NO	YES	NO	YES	NO	YES	NO
6	YES	NO	YES	NO	YES	NO	YES	NO
7	YES	NO	YES	NO	YES	NO	YES	NO
8	YES	NO	YES	NO	YES	NO	YES	NO
9	YES	NO	YES	NO	YES	NO	YES	NO
10	YES	NO	YES	NO	YES	NO	YES	NO
11	YES	NO	YES	NO	YES	NO	YES	NO
12	YES	NO	YES	NO	YES	NO	YES	NO
13	YES	NO	YES	NO	YES	NO	YES	NO
14	YES	NO	YES	NO	YES	NO	YES	NO
15	YES	NO	YES	NO	YES	NO	YES	NO
16	YES	NO	YES	NO	YES	NO	YES	NO
17	YES	NO	YES	NO	YES	NO	YES	NO
18	YES	NO	YES	NO	YES	NO	YES	NO
19	YES	NO	YES	NO	YES	NO	YES	NO
20	YES	NO	YES	NO	YES	NO	YES	NO
21	YES	NO	YES	NO	YES	NO	YES	NO
22	YES	NO	YES	NO	YES	NO	YES	NO
23	YES	NO	YES	NO	YES	NO	YES	NO
24	YES	NO	YES	NO	YES	NO	YES	NO
25	YES	NO	YES	NO	YES	NO	YES	NO
26	YES	NO	YES	NO	YES	NO	YES	NO
27	YES	NO	YES	NO	YES	NO	YES	NO
28	YES	NO	YES	NO	YES	NO	YES	NO
29	YES	NO	YES	NO	YES	NO	YES	NO
30	YES	NO	YES	NO	YES	NO	YES	NO
Totals	26	4	28	2	18	12	22	8
% Intervals on task	86.6%		93.3%		60.0%		73.3%	

Yes = On-task No = Off-task

Figure 4.6 Observation form used for whole-interval recording of on-task behavior by four students during independent seatwork time.

Adapted from *Smiley Faces and Spinners: Effects of Self-Monitoring of Productivity with an Indiscriminable Contingency of Reinforcement on the On-Task Behavior and Academic Productivity by Kindergarteners During Independent Seatwork* by R. L. Ludwig, 2004, p. 101. Unpublished master's thesis, The Ohio State University. Used by permission.

Partial-Interval Recording

When using **partial-interval recording**, the observer records whether the behavior occurred *at any time* during the interval. Partial-interval time sampling is not concerned with how many times the behavior occurred during the interval or how long the behavior was present, just that it occurred at some point during the interval. If the target behavior occurs multiple times during the interval, it is still scored as occurring only once. An observer using partial-interval recording to measure a student's disruptive behavior would mark an interval if disruptive behavior of any form meeting the target behavior definition occurred for any amount of time during the interval. That is, an interval would be scored as disruptive behavior even if the student were disruptive for only 1 second of a 6-second interval. Because of this, data obtained with partial-interval recording often overestimate the overall percentage of the observation period (i.e., total duration) that the behavior actually occurred.

Partial-interval data, like whole-interval data, are most often reported as a percentage of total intervals in which the target behavior was scored. Partial-interval data are used to represent the proportion of the entire observation period in which the target behavior occurred, but the results of partial-interval recording, unlike those of whole-interval recording, do not provide any information on duration per occurrence. That is because any instance of the target behavior, regardless of how brief its duration, will cause an interval to be scored.

If the observer uses partial-interval recording with brief observation intervals to measure discrete responses of short duration per occurrence, the data obtained provide a crude estimate of the minimum rate of responding. For example, data showing that a behavior measured by partial-interval recording consisting of 6-second contiguous intervals (i.e., successive intervals are not separated by time in which the behavior is not observed) occurred in 50% of the total intervals indicate a minimum response rate of five responses per minute (on the average, at least one response occurred in 5 of the 10 intervals per minute). Although partial-interval recording often overestimates the total duration, it is likely to underestimate the rate of a high-count behavior. This is because an interval in which a person made eight nonverbal sounds would be scored the

same as an interval in which the person made only one sound. When the evaluation and understanding of a target behavior require a more sensitive measure than interval recording, use rate of response.

Because an observer using partial-interval recording needs to record only that a behavior has occurred at any point during each interval (compared to having to watch the behavior throughout the entire interval with whole-interval recording), it is possible to measure multiple behaviors concurrently. Figure 4.7 shows a portion of a form for measuring four response classes by three students, using partial-interval recording with 20-second intervals. The observer watches Student 1 throughout the first 20-second interval, Student 2 for the next 20 seconds, and Student 3 for the next 20 seconds. Each student is observed for 20 seconds out of each minute of the observation period. If a student engages in any of the behaviors being measured at any time during an observation interval, the observer marks the letter(s) corresponding to those behaviors. If a student engages in none of the behaviors being measured during an interval, the observer marks N to indicate no occurrences of the target behaviors. For example, during the first interval in which Student 1 was observed, he said, "Pacific Ocean" (an academic response). During the first interval in which Student 2 was observed, she left her seat and threw a pencil (a behavior within the response class "other disruptive behavior"). Student 3 emitted none of the four target behaviors during the first interval that he was observed.

Momentary Time Sampling

An observer using **momentary time sampling** records whether the target behavior occurred *at the moment* the time interval ends. If conducting momentary time sampling with 1-minute intervals, an observer would look at the person at the 1-minute mark of the observation period, determine immediately whether the target behavior was occurring, and indicate that decision on the recording form. One minute later (i.e., 2 minutes into the observation period), the observer would look again at the person and then score the presence or absence of the target behavior. This procedure would continue until the end of the observation period.

As with interval recording methods, data from momentary time sampling are typically reported as percentages of the total intervals in which the behavior occurred and are used to estimate the proportion of the total observation period that the behavior occurred.

A major advantage of momentary time sampling is that the observer does not have to attend continuously to measurement, whereas interval-recording methods demand the undivided attention of the observer.

Because the person is observed for only a brief moment, much behavior will be missed with momentary time sampling. Momentary time sampling is used primarily to measure continuous activity behaviors such as engagement with a task or activity, because such behaviors are easily identified. Momentary time sampling is not recommended for measuring low-count, short-duration behaviors.

Numerous studies have compared measures obtained by momentary time sampling and interval recording using intervals of varying duration with measures obtained by continuous duration recording (e.g., Alvero, Struss, & Rappaport, 2007; Gunter, Venn, Patrick, Miller, & Kelly, 2003; Hanley, Cammilleri, Tiger, & Ingvarsson, 2007; Meany-Daboul, Roscoe, Bourret, & Ahearn, 2007; Powell, Martindale, Kulp, Martindale, & Bauman, 1977; Saudargas & Zanolli 1990; Test & Heward, 1984). In general, this research has found that momentary time sampling both overestimates and underestimates continuous duration measures when intervals are greater than 2 minutes. With intervals up to 1 minute, the data obtained using momentary time sampling produce data paths similar to those of continuous duration recording.

Planned Activity Check

A variation of momentary time sampling, **planned activity check (PLACHECK)**, uses head counts to measure "group behavior." A teacher using PLACHECK observes a group of students at the end of each time interval, counts the number of students engaged in the targeted activity, and records the tally with the total number of students in the group. Doke and Risley (1972) used data obtained by PLACHECK measurement to compare group participation in required and optional

	1	2	3	4
Student 1	A T S D N	A T S D N	A T S D N	A T S D N
Student 2	A T S D N	A T S D N	A T S D N	A T S D N
Student 3	A T S D N	A T S D N	A T S D N	A T S D N

Key:
A = Academic response
T = Talk-out
S = Out-of-seat
D = Other disruptive behavior
N = No occurrences of target behaviors

Figure 4.7 Portion of a form used for partial-interval recording of four response classes by three students.

before-school activities. Observers tallied the number of students in either the required or the optional activity area at the end of 3-minute intervals, and then the number of children actually participating in an activity in either area. They reported these data as separate percentages of children participating in required or optional activities.

Dyer, Schwartz, and Luce (1984) used a variation of PLACHECK to measure the percentage of students with disabilities living in a residential facility who were engaged in age-appropriate and functional activities. As students entered the observation area, they were observed individually for as long as it took to determine the activity in which they were engaged. The students were observed in a predetermined order not exceeding 10 seconds per student.

In a study examining the effects of response cards on the disruptive behavior of third-graders during daily math lessons, Armendariz and Umbreit (1999) recorded at 1-minute intervals whether each student in the class was being disruptive. By combining the PLACHECK data obtained across all of the no-response cards (baseline) sessions, and graphing those results as the percentage of students who were disruptive at each 1-minute mark, and doing the same with the PLACHEK data from all response cards sessions, Armendariz and Umbreit created a clear and powerful picture of the differences in "group behavior" from the beginning to the end of a typical lesson in which response cards were or were not used.

Other variations of the PLACHECK measurement can be found in the literature, though they usually are called time sampling or momentary time sampling. For example, McKenzie and colleagues have developed two PLACHECK-type observational systems, SOPLAY and SOPARC, for measuring children's physical activity in school and community play environments (McKenzie & Cohen, 2006; McKenzie, Marshall, Sallis, & Conway, 2000). "SOPLAY and SOPARC use a group momentary time-sampling format (i.e., serial observation 'snapshots') to record the physical activity level (i.e., sedentary, walking/moderate, vigorous) in specific target areas. . . . Counting the number of people in an area is itself an important endeavor because it provides an indication of how well the space/facility supports school or community physical activity goals" (McKenzie, 2016, p. 335, 336).

Recognizing Discrepancies Among and Between Time Sampling Measures and Continuous Measurement

As stated previously, all time sampling methods provide only an estimate of the actual occurrence of the behavior. Different time sampling procedures produce different results, which can influence decisions and interpretations. A series of studies by Rapp and colleagues compared the sensitivity of momentary time sampling (MTS) and partial-interval recording (PIR) to detect small, moderate, and large behavior changes evident with continuous duration measures (Carroll, Rapp, Colby-Dirksen, & Lindenberg, 2009; Devine, Rapp, Testa, Henrickson, & Schnerch, 2011; Rapp, Colby-Dirksen, Michalski, Carroll, & Lindenberg, 2008). These researchers compared measures obtained by MTS and PIR of various interval sizes (10-sec, 20-sec, 30-sec, 1-min, and 2-min intervals) within varied observation periods (10-min, 30-min, and 60-min sessions) with data provided by continuous measurement (duration or event recording). Rapp et al. (2008) found that MTS with intervals up to 30 seconds reliably detected most moderate and large changes in duration events, that MTS with 1-minute intervals detected most large changes in duration events, and that PIR with 10-second or larger intervals failed to detect a large percentage of changes in duration events. Given longer observation sessions, momentary time sampling with interval sizes up to 30 sec detected a wide range of changes in both duration and rate events.

Devine and colleagues (2011) found that momentary time sampling with interval sizes up to 30 seconds detected a wide range of changes in duration and rate events during lengthier observation periods. These authors concluded that

> [P]ractitioners with limited resources may measure behavior changes with smaller interval sizes of PIR or MTS during shorter observation sessions or with a larger interval size of MTS during lengthier sessions with comparable sensitivity. For instance, 10-s MTS during 10-min sessions and 30-s MTS during 30-min sessions yielded comparable detection of changes in duration events. Similarly, both 10-s PIR during 10-min sessions and 30-s MTS during 30-min sessions were sensitive to changes in frequency events. In both cases, the number of observations required in the 10-min session is the same as the number of observations required in the 30-min session. (pp. 120–121)

Figure 4.8 illustrates just how different the results obtained by measuring the same behavior with different time sampling methods can be. The shaded bars indicate when the behavior was occurring within an observation period divided into 10 contiguous intervals. The shaded bars reveal the three-dimensional quantities of behavior: *repeatability* (seven instances of the behavior), *temporal extent* (the duration of each response) and *temporal locus* (IRT time is depicted by the space between the shaded bars).

Because the time sampling methods used in applied behavior analysis are most often viewed and interpreted as measures of the proportion of the total observation period in which the behavior occurred, it is important to compare the results of time sampling methods with those obtained by continuous measurement of duration. Continuous measurement reveals that the behavior depicted in Figure 4.8 occurred 55% of the time during the observation period. When the same behavior during the same observation period was recorded using whole-interval recordings, the measure obtained grossly underestimated the actual occurrence of the behavior (i.e., 30% versus 55%), partial-interval recording grossly overestimated the actual occurrence (i.e., 70% versus 55%), and momentary time sampling yielded a fairly close estimate of actual occurrence of the behavior (50% versus 55%).

Although momentary time sampling resulted in a measure that most closely approximated the actual behavior, it is not always the preferred method. Different distributions of the

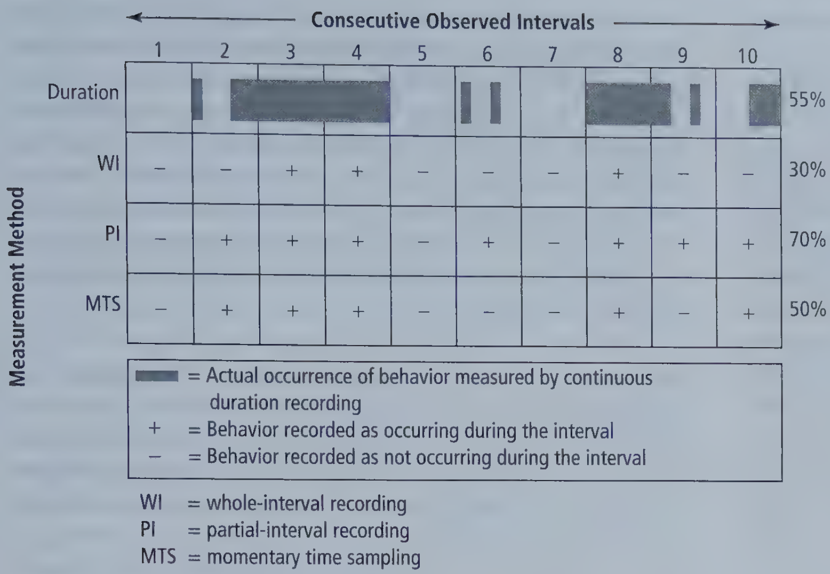


Figure 4.8 Comparing measures of the same behavior obtained by three different time sampling methods with measure obtained by continuous duration recording.

behavior (i.e., temporal locus) during the observation period, even at the same overall count and duration as the session shown in Figure 4.8, would result in widely different outcomes from each of the three time sampling methods.

Discrepancies between measures obtained by different measurement methods are usually described in terms of the relative accuracy or inaccuracy of each method. However, accuracy is not the issue here. If the shaded bars in Figure 4.8 represent the true value of the behavior, then each of the time sampling methods was conducted with complete accuracy and the resulting data are what should be obtained by applying each method. An example of the inaccurate use of one of the measurement methods would be if the observer using whole-interval recording had marked the behavior as occurring in Interval 2, when the behavior did not occur according to the rules of whole-interval recording.

But if the behavior actually occurred for 55% of the observation period, what should we call the wrong and misleading measures of 30% and 70% if not inaccurate? In this case, the misleading data are artifacts of the measurement methods used to obtain them. An **artifact** is a phenomenon that appears to exist because of the way it is examined or measured. The 30% measure obtained by whole-interval recording and the 70% measure obtained by partial-interval recording are artifacts of the way those measures were conducted. The fact that data obtained from whole-interval and partial-interval recording consistently underestimate and overestimate, respectively, actual occurrence of behavior as measured by continuous duration recording is an example of well-known artifacts.

It is clear that interval measurement and momentary time sampling result in some variability in the data analysis, which must be considered carefully when interpreting results obtained with these measurement methods. Some common causes of measurement artifacts and how to avoid them are discussed in Chapter 5.

MEASURING BEHAVIOR BY PERMANENT PRODUCTS

Behavior can be measured in real time by observing a person's actions and recording responses of interest as they occur. For example, a teacher can tally the number of times a student raises her hand during a class discussion. Some behaviors can be measured in real time by recording their effects on the environment as those effects are produced. For example, a baseball batting coach advances a handheld counter each time a batter hits a pitch to the right-field side of second base.

Some behaviors can be measured after they have occurred. A behavior that produces consistent effects on the environment can be measured after it has occurred if the effects, or products left behind by the behavior, remain unaltered until measurement is conducted. For example, if the flight of the baseballs hit by a batter during a practice session were not impeded, and the balls were left lying on the ground, a hitting instructor could collect data on the batter's performance after the batter's turn was completed by counting each ball found lying in fair territory on the right-field side of second base.

Measuring behavior after it has occurred by observing the effects the behavior produced on the environment is known as **measurement by permanent product**. A permanent product is a change in the environment produced by a behavior that lasts long enough for measurement to take place.

Although often described erroneously as a method for measuring behavior, measurement by permanent product does not refer to any particular measurement procedure or method. Instead, measurement by permanent product refers to the time of measurement (i.e., after the behavior has occurred) and the medium (i.e., the effect of the behavior, not the behavior itself) by which the measurer comes in contact with (i.e., observes) the behavior. All of the methods for measuring behavior described

in this chapter—event recording, timing, and time sampling—can be used to measure permanent products.

Permanent products can be natural or contrived outcomes of the behavior. Permanent products are natural and important outcomes of a wide range of socially significant behaviors in educational, vocational, domestic, and community environments. Examples in education include compositions written (Dorow & Boyle, 1998), computations of math problems written (Skinner, Fletcher, Wildmon, & Belfiore, 1996), spelling words written (McGuffin, Martz, & Heron, 1997), worksheets completed (Alber, Heward, & Hippler, 1999), homework assignments turned in (Alber, Nelson, & Brennan, 2002), and test questions answered (e.g., Gardner, Heward, & Grossi, 1994). Behaviors such as mopping floors and dishwashing (Grossi & Heward, 1998), incontinence (Adkins & Matthews, 1997), drawing bathroom graffiti (Mueller, Moore, Doggett, & Tingstrom, 2000), recycling (Brothers, Krantz, & McClannahan, 1994), stealing food (Maglieri, DeLeon, Rodriguez-Catter, & Sevin, 2000), and picking up litter (Powers, Osborne, & Anderson, 1973) can also be measured by the natural and important changes they make on the environment. Carbon monoxide (CO) levels in expired breath rise with increasing cigarette consumption and decline with abstention. Romanowich and Lamb's (2015) smoking abstinence study measured CO in exhaled breath as a permanent product measure. Participants took a deep breath, held it for 20 seconds, and then exhaled over 20 seconds into the mouthpiece of a Vitalograph™ CO monitor.

Many socially significant behaviors have no direct effects on the physical environment. Reading orally, sitting with good posture, and repetitive hand flapping leave no natural products in typical environments. Nevertheless, measurement of such behaviors can often be accomplished using contrived permanent products. For example, by audio recording students as they read out loud (Eckert, Ardoin, Daly, & Martens, 2002), video recording a girl sitting in class (Schwarz & Hawkins, 1970), and video recording a boy flapping his hands (Ahearn, Clark, Gardenier, Chung, & Dube, 2003), researchers obtained contrived permanent products for measuring these behaviors.

Contrived permanent products are sometimes useful in measuring behaviors that have natural permanent products that are only temporary. For example, Goetz and Baer (1973) measured variations in the form of children's block building from photographs taken of the constructions made by the children, and Twohig and Woods (2001) measured the length of fingernails from photographs of nail-biters' hands.

Advantages of Measuring Permanent Products

Frees Practitioner for Other Tasks. Not having to observe and record behavior as it occurs enables the practitioner to attend to other tasks during the observation period. For example, a teacher who audio records students' questions, comments, and talk-outs during a class discussion can concentrate on what her students are saying, provide individual help, and so on.

Enables Measurement of Some Behaviors That Occur at Inconvenient or Inaccessible Times and Places. Many socially significant behaviors occur at times and places that are inconvenient or inaccessible to the researcher or practitioner. Measuring permanent products is indicated when observing the target behavior as it occurs would be difficult because the behavior occurs infrequently, in various environments, or for extended periods. For example, a music teacher can have a guitar student make audio recordings of portions of daily practice sessions at home.

Measurement May Be More Accurate, Complete, and Continuous. Although measuring behavior as it occurs provides the most immediate access to data, it does not necessarily yield the most accurate, complete, and representative data. An observer measuring behavior from permanent products can take his time, rescore the worksheet, or listen to and watch the video recording again. Video recording enables the observer to slow down, pause, and repeat portions of the session—to literally "hold still" the behavior so it can be examined and measured repeatedly if necessary. The observer might see or hear additional nuances and aspects of the behavior, or other behaviors that he overlooked or missed altogether during live performances.

Video- or audio-recording behavior enables continuous measurement of all instances of a target behavior. Having this permanent product of all instances lends itself to later scoring by using the built-in calibrated digital timer (e.g., on a VCR or digital recorder), set to zero seconds (or the first frame) at the beginning of the session, to note the exact time of behavior onset and offset. Further, software programs facilitate data collection and analysis based on exact timings. PROCODER is a software system for facilitating the collection and analysis of video-recorded behavior. According to Miltenberger, Rapp, and Long (1999), "With a recording of the exact time of the onset and offset of the target behavior in the observation session, we are able to report the frequency (or rate) or the duration of the behavior" (p. 119).[8]

Measurement by permanent product enables data collection on more participants. An observer can view a video recording once and measure one participant's behavior, then replay the recording and measure a second participant's behavior.

Facilitates Data Collection for Interobserver Agreement and Treatment Integrity. Video- or audio recording assists with data collection tasks such as obtaining interobserver agreement data (see Chapter 5) and assessing treatment integrity (Chapter 10). Permanent products make possible the repeated measurement of behavior, eliminating the need to bring multiple observers into the research or treatment setting.

Enables Measurement of Complex Behaviors and Multiple Response Classes. Permanent products, especially video recordings of behavior, enable measurement of complex behaviors and

multiple response classes in busy social environments. Schwarz and Hawkins (1970) obtained measures of the posture, voice volume, and face touching of an elementary student from videotapes taken during two class periods. The three behaviors were targeted as the result of operationalizing the girl's "poor self-esteem." The researchers were able to view the tapes repeatedly and score them for different behaviors. In this study, the girl also watched and evaluated her behavior on the videotapes as part of the intervention.

Figure 4.9 shows an example of a recording form used by Silvestri (2004) to measure three types of statements by teachers—generic positive, specific positive, and negative—from audiotapes of classroom lessons. (See Figure 3.11 for definitions of these behaviors.) The movement, multiple voices, and general commotion that characterize any classroom would combine to make it very difficult, if not impossible, for a live observer to consistently detect and accurately record these behaviors. Each teacher participant in the study wore a small wireless microphone that transmitted a signal to a receiver connected to a cassette tape recorder.

Determining Whether Measurement by Permanent Product Is Appropriate

The advantages of measurement by permanent product are considerable, and it may seem as though permanent product measurement is always preferable to real-time measurement. Answering four questions will help practitioners and researchers determine whether measurement by permanent product is appropriate: Is real-time measurement needed? Can the behavior be measured by permanent product? Will obtaining a contrived permanent product unduly affect the behavior? How much will it cost?

Participant: T1 Date of session: 4/23 Exp. condition: Self-scoring generalization

Observer: Susan Date of observation 4/23 Duration of observation: 15:00 Decimal: 15.0

Transcribe positive and negative statements, corresponding time indexes, and time indexes of repeated statements in the boxes below.

Generic Positive (transcribe first instance)	Time index	Time index of repeats		Specific Positive (transcribe first instance)	Time index	Time index of repeats		Negative (transcribe first instance)	Time index	Time index of repeats	
Excellent	0:17	1:37	3:36	I like how you helped her.	1:05						
		4:00	4:15								
		7:45	9:11	Thank you for not talking.	1:57	2:10	3:28				
		10:22	10:34								
Beautiful	0:26	1:44	1:59	Good—nice big word.	2:45	6:53	8:21				
		9:01	11:52			9:56					
		13:09		You raised your hand, beautiful job.	3:37	4:33					
Good	0:56	1:22	4:42	Good job, that's a new word.	3:46						
		5:27	5:47								
		6:16	6:38	Thank you for paying attention.	4:56						
		8:44	9:25								
Smart	5:14	7:06	11:59	Thank you for not writing.	7:50						
High five	8:00										
Count: 28	% repeats: 83%			Count: 13	% repeats 46%			0	% repeats:		
No. per min. 1.9				No. per min. 0.9							
All Positives:	Count: 41			No. per min. 2.7	% repeats 71%						

Figure 4.9 Data collection form for recording count and temporal locus of three classes of teacher statements from video-recorded sessions.

From *The Effects of Self-Scoring on Teachers' Positive Statements During Classroom Instruction* by S. M. Silvestri (2004), p. 124. Unpublished doctoral dissertation, The Ohio State University. Used by permission.

Is Real-Time Measurement Needed?

A defining feature of applied behavior analysis and one of its major strengths is using data-based decisions concerning treatment procedures and experimental conditions. Data-based decision making requires more than direct and frequent measurement of behavior; it also requires ongoing and timely access to the data those measures provide. Measuring behavior as it occurs provides the most immediate access to the data. Although real-time measurement by permanent product can be conducted in some situations (e.g., counting each batted ball that lands to the right of second base as a hitter takes batting practice), measurement by permanent product will most often be conducted after the instructional or experimental session has ended.

Measures taken from video or audio recordings cannot be obtained until the recordings are viewed after a session has ended. If treatment decisions are made on a session-by-session basis, this behavior-to-measurement delay poses no problem as long as the data can be obtained from the recordings prior to the next session. However, when moment-to-moment treatment decisions must be made according to the participant's behavior during the session, real-time measurement is necessary. Consider a behavior analyst trying to reduce the rate of a person's self-injurious behavior (SIB) by providing access to a preferred stimulus contingent on increasing durations of time without SIB. Accurate implementation of the treatment protocol would require the real-time measurement of IRTs.

Can the Behavior Be Measured by Permanent Product?

Not all behaviors are suitable for measurement by permanent product. Some behaviors affect fairly permanent changes in the environment that are not reliable for the purposes of measurement. For example, SIB often produces long-lasting effects (bruises, welts, and even torn and bleeding skin) that could be measured after the occurrence of the behavior. But accurate measures of SIB could not be obtained by examining the client's body on a regular basis. The presence of discolored skin, abrasions, and other such marks would indicate that the person had been injured, but many important questions would be left unanswered. How many times did SIB occur? Were there acts of SIB that did not leave observable marks on the skin? Was each instance of tissue damage the result of SIB? These are important questions for evaluating the effects of any treatment. Yet, they could not be answered with certainty because these permanent products are not precise enough for the measurement of the SIB. Behaviors suitable for measurement with permanent product must meet two rules.

Rule 1: Each occurrence of the target behavior must produce the same permanent product. The permanent product must be a result of every instance of the behavior measured. All topographical variants of the target behavior and all responses of varying magnitude that meet the definition of target behavior must produce the same permanent product. Measuring an employee's work productivity by counting the number of correctly assembled widgets in his "completed work" bin conforms to this rule. An occurrence of the target behavior in this case is defined functionally as a correctly assembled widget. Measuring SIB by marks on the skin does not meet Rule 1 because some SIB responses will not leave discernible marks.

Rule 2: The permanent product can be produced only by the target behavior. This rule requires that the permanent product cannot result from (a) any behaviors by the participant other than the target behavior or (b) the behavior of any person other than the participant. Using the number of correctly assembled objects in the employee's "completed work" bin to measure his productivity conforms to Rule 2 *if* the observer can be assured that (a) the employee put no objects in his bin that he did not assemble and (b) none of the assembled objects in the employee's bin were put there by anyone other than the employee. Using marks on the skin as a permanent product for measuring self-injurious behavior also fails to meet Rule 2. Marks on the skin could be produced by other behaviors by the person (e.g., running too fast, which caused him to trip and hit his head; stepping in poison ivy) or by the behavior of other people (e.g., struck by another person).

Will Obtaining a Contrived Permanent Product Unduly Affect the Behavior?

Practitioners and researchers should always consider reactivity—the effects of the measurement procedure on the behavior being measured. Reactivity is most likely when observation and measurement procedures are obtrusive. Obtrusive measurement alters the environment, which may in turn affect the behavior being measured. Permanent products obtained using recording equipment, the presence of which may cause a person to behave differently, are called contrived permanent products. For example, recording conversations might encourage the participant to talk less or more. But it should be recognized that reactivity to the presence of human observers is a common phenomenon, and reactive effects are usually temporary (e.g., Haynes & Horn, 1982; Kazdin, 1982, 2001). Even so, it is appropriate to anticipate the influence that equipment may have on the target behavior.

How Much Will It Cost to Obtain and Measure the Permanent Product?

The final issues to address in determining the appropriateness of measuring a target behavior by permanent product are availability, cost, and effort. If recording equipment is required to obtain a contrived product of the behavior, is it currently available? If not, how much will it cost to buy or rent the equipment? How much time will it take to learn to use the equipment initially? How difficult and time-consuming will it be to set up, store, and use the equipment for the duration of the study or behavior change program?

MEASUREMENT TOOLS

Low-Tech

The low-tech measurement tools described here have a long history in applied behavior analysis. Researchers and practitioners will continue to use these "old school" tools described here because they are functional, easy to use, and inexpensive.

- *Pencil-and-paper data sheets.* Observers record occurrences and nonoccurrences of behavior on data sheets such as those shown in Figures 4.5, 4.6, 4.7, and 4.9.
- *Wrist counters.* Wrist counters are useful for tallying behavior. Most wrist counters record from 0 to 99 responses. These counters can be purchased from sporting goods stores or large department stores.
- *Hand-tally digital counters.* Similar to wrist counters, hand-tally digital counters are frequently used in grocery chain stores, cafeterias, military mess halls, and tollgates to tally the number of people served. These mechanical counters are available in single or multiple channels and fit comfortably in the palm of the hand. With practice, applied behavior analysts can operate the multiple-channel counters rapidly and reliably with just one hand. Digital counters can be obtained from office supply stores.
- *Abacus wrist and shoestring counters.* Landry and McGreevy (1984) described two types of abacus counters for measuring behavior. The abacus wrist counter is made from pipe cleaners and beads attached to a leather wristband to form an abacus with rows designated as ones and tens. An observer can tally from 1 to 99 occurrences of behavior by sliding the beads in abacus fashion. Responses are tallied in the same way on an abacus shoestring counter except the beads slide on shoestrings attached to a key ring, which is attached to the observer's belt, belt loop, or some other piece of clothing, such as a buttonhole.
- *Masking tape.* Practitioners tally behavioral events on strips of masking tape attached to their wrists, clothing, or desk.
- *Pennies, buttons, paper clips.* Observer moves one item from one pocket to another each time the target behavior occurs.

High-Tech

Automated measurement of subjects' responses has been standard practice in the experimental analysis of behavior (EAB) since Skinner's invention of the cumulative recorder (Lattal, 2004; Sakagami & Lattal 2016). In today's EAB laboratory, the electro-mechanical cumulative recorder has been replaced by computer-based measurement. Researchers can program sophisticated hardware and software systems to continuously measure rate, duration, latency, IRTs, and/or magnitude on multiple response classes throughout experimental sessions.

Computer-assisted measurement and data analysis tools have become increasingly sophisticated and useful for applied behavior analysts. Developers have produced data collection and analysis software for observational measurement using laptops (Bullock, Fisher, & Hagopian, 2017), mobile devices such as tablets and smartphones (Operant Systems Inc, 2017), personal digital assistants (PDAs) (Fogel, Miltenberger, Graves, & Koehler, 2010), and wearables that automatically record movement such as steps taken and distance traveled (Hayes & Van Camp, 2015; Valbuena, Miltenberger, & Solley, 2015).

Practitioners using apps designed for data collection during treatment sessions tap icons on a touchscreen representing response classes, prompt levels, and consequences (e.g., Richard, Taylor, DeQuinzio, & Katz, 2010; Sleeper et al., 2017). At session's end, practitioners can add notes and tap an icon to produce graphs immediately available for inspection.

Developments in microchip technology have advanced the measurement and data analysis capabilities of these systems, and have made the software increasingly easier to learn and apply. Observers using some of these computerized tools can record multiple frequency- and duration-based behaviors, including discrete trials, the number of responses per unit of time, duration, latency, interresponse time (IRT), and fixed and variable intervals for time sampling measurement. These systems can present data calculated as rate, duration, latency, IRT, percentage of intervals, percentage of trials, and conditional probabilities (e.g., BDataPro [Bullock, Fisher, and Hagopian, 2017]; MOOSES [Tapp, Wehby, & Ellis, 1995]; BEST—Behavioral Evaluation Strategy and Taxonomy—Sidenar, Shabani, & Carr, 2004).[9]

Computer-based (automated) observation and measurement systems have distinct advantages over non-automated measurement when working with aggregates of rates of response, time-series analyses, conditional probabilities, sequential dependencies, interrelationships, and combinations of events. Automated measurement facilitates the clustering and analyzing of complex data sets. Because these systems allow for the simultaneous recording of multiple behaviors across multiple dimensions, outputs can be examined and analyzed from perspectives that would be difficult and time-consuming with paper-and-pencil methods.

In addition to recording and calculating data, some systems produce graphs (e.g., Sleeper et al., 2017) and analyses of interobserver agreement (IOA) (e.g., smaller/larger, overall, occurrence, nonoccurrence) and measurement from audio and video documents. The automated computer-driven systems, as compared to the commonly used paper-and-pencil data-recording and analysis methods, have the potential to improve IRAs, the reliability of observational measurements, and the efficiency of data calculation (Kahng & Iwata, 1998). Automated measurement has the capacity to eliminate many human observer errors (e.g., observer drift, expectancy bias).

We believe continued advances in digital measurement tools will further enrich research and practice in applied behavior analysis. Some cautions, however, are associated with automated measurement of behavior. The equipment must be monitored because machines can break, and programmers can make mistakes. Kahng, Ingvarsson, Quigg, Seckinger, and Teichman (2011) caution applied behavior analysts' use of automated measurement. "Computer programs of this kind should be used only if they provide improvements in data collection that are of sufficient importance considering the behavior of interest and the ultimate goals of data collection." (p. 117).

SELECTING A MEASUREMENT METHOD

Although it may go without saying that the measurement method must fit the target behavior dimension one wants to measure (e.g., event recording for rate, timing for duration and latency), selecting the most appropriate measurement method for a given situation is hardly as simple as that. Researchers and practitioners must also consider the behavior change goals and expected direction of behavior change, the relative ease of detecting occurrences of the behavior, the environments where and times when the behavior will be measured, and the availability and skills of personnel who will observe and record the

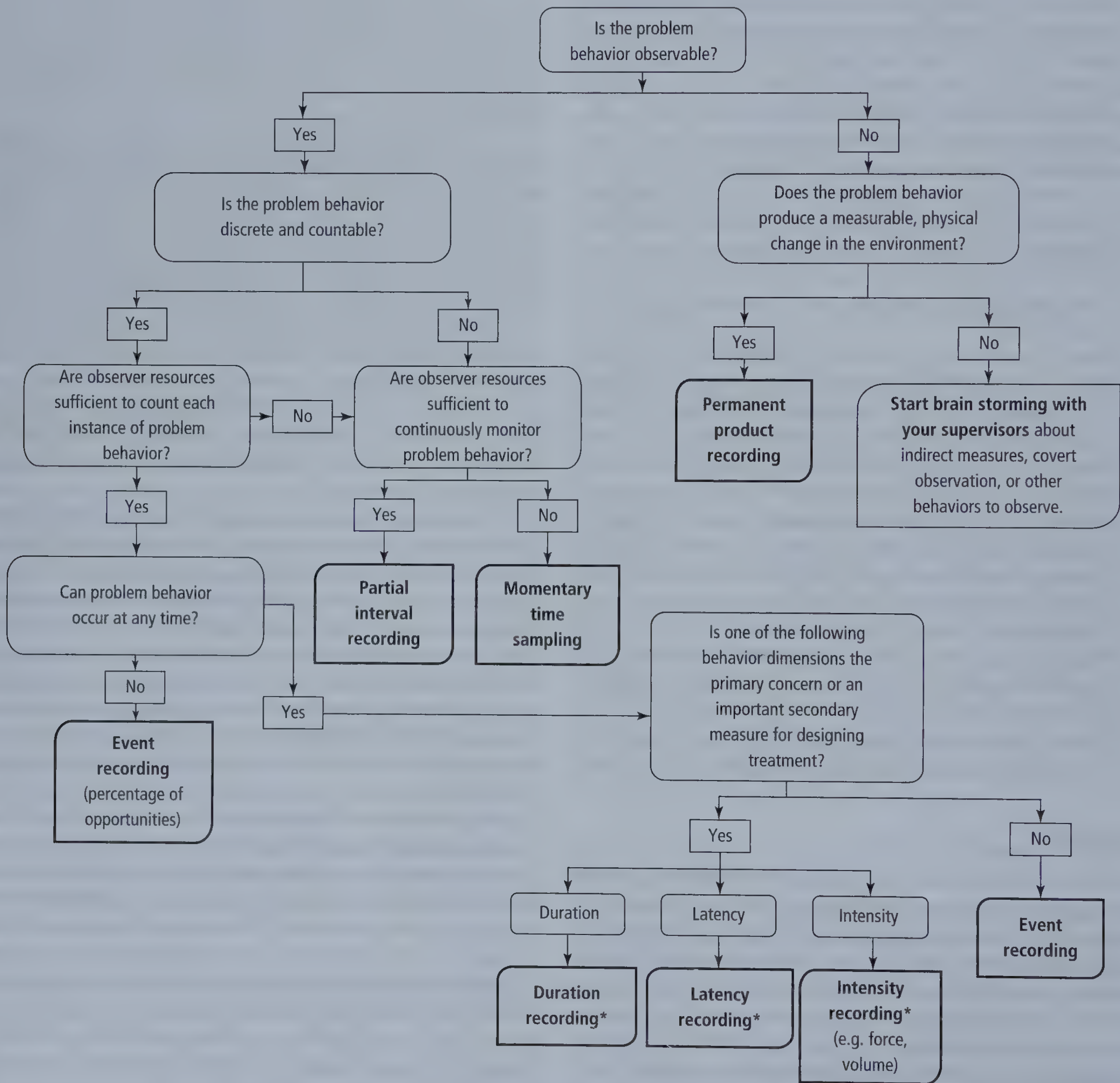


Figure 4.10 Decision-making model for selecting measurement procedures for problem behavior. *Note*: Asterisks denote measurement procedures that also generate count data.

From "A Proposed Model for Selecting Measurement Procedures for the Assessment and Treatment of Problem Behavior" by L. A. LeBlanc, P. B. Raetz, T. P. Sellers, and J. E. Carr, 2016, *Behavior Analysis in Practice*, 9(1), p. 79. Copyright 2016 by the Association for Behavior Analysis International. Used by permission.

behavior (Fiske & Delmolino, 2012; Gast 2014; LeBlanc, Raetz, Sellers, & Carr, 2016). We believe the decision-making model selecting measurement methods proposed by LeBlanc and colleagues (2016) provides useful guidance to practitioners (see Figure 4.10).

Practitioners should balance the resources required to observe and measure the target behavior with the planned uses of the data. The old adage "Just because you can, doesn't mean you should" can be a useful guide when designing behavioral measurement systems. Recording every response emitted by a client and every therapist-delivered prompt and consequence is not always necessary (Lerman, Dittlinger, Fentress, & Lanagan, 2011; Najdowski et al., 2009). Collect the type and number of behavioral measures needed to make timely, data-based decisions about treatment effects. Measurement that does not provide a decision maker with close, continuous contact with relevant outcome data may be superfluous (cf. Bushell & Baer, 1994).

Additional considerations and issues with respect to designing behavioral measurement systems are discussed in Chapter 5.

SUMMARY

Definition and Functions of Measurement in Applied Behavior Analysis

1. Measurement is the process of applying quantitative labels to observed properties of events using a standard set of rules.
2. Measurement is how scientists operationalize empiricism.
3. Without measurement, all three levels of scientific knowledge—description, prediction, and control—would be relegated to guesswork and subjective opinions.
4. Applied behavior analysts measure behavior to obtain answers to questions about the existence and nature of functional relations between socially significant behavior and environmental variables.
5. Practitioners measure behavior before and after treatment to evaluate the overall effects of interventions (summative evaluation) and frequent measures of behavior during treatment (formative assessment) to guide decisions concerning the continuation, modification, or termination of treatment.
6. Without frequent measures of the behavior targeted for intervention, practitioners may (a) continue an ineffective treatment when no real behavior change occurred or (b) discontinue an effective treatment because subjective judgment detects no improvement.
7. Measurement also helps practitioners optimize their effectiveness; verify the legitimacy of practices touted as "evidence based"; identify treatments based on pseudoscience, fad, fashion, or ideology; be accountable to clients, consumers, employers, and society; and achieve ethical standards.

Measurable Dimensions of Behavior

8. Because behavior occurs within and across time, it has three dimensional quantities: repeatability (i.e., count), temporal extent (i.e., duration), and temporal locus (i.e., when behavior occurs). These properties, alone and in combination, provide the basic and derivative measures used by applied behavior analysts.
9. Count is a simple tally of the number of occurrences of a behavior.
10. Rate is a ratio of count per observation period; it is expressed as count per standard unit of time.
11. Celeration is a measure of the change (acceleration or deceleration) in rate of responding per unit of time.
12. Duration is the amount of time from the onset to the end point of a behavior.
13. Response latency is a measure of the elapsed time between the onset of a stimulus and the initiation of a subsequent response.
14. Interresponse time (IRT) is the amount of time that elapses between two consecutive instances of a response class.
15. Percentage, a ratio formed by combining the same dimensional quantities, expresses the proportional quantity of an event in terms of the number of times the event occurred per 100 opportunities that the event could have occurred.
16. Trials-to-criterion is a measure of the number of response opportunities needed to achieve a predetermined level of performance.
17. Although form (i.e., topography) and intensity of responding (i.e., magnitude) are not fundamental dimensional quantities of behavior, they are important quantitative parameters for defining and verifying the occurrence of many response classes.
18. Topography refers to the physical form or shape of a behavior.
19. Magnitude refers to the force or intensity with which a response is emitted.

Methods for Measuring Behavior

20. Event recording encompasses a wide variety of procedures for detecting and recording the number of times a behavior of interest is observed.

21. Time sampling refers to a variety of methods for observing and recording behavior during intervals or at specific moments in time.

22. Observers using whole-interval recording divide the observation period into a series of equal time intervals. At the end of each interval, they record whether the target behavior occurred throughout the entire interval.

23. Observers using partial-interval recording divide the observation period into a series of equal time intervals. At the end of each interval, they record whether behavior occurred at any point during the interval.

24. Observers using momentary time sampling divide the observation period into a series of time intervals. At the end of each interval, they record whether the target behavior is occurring at that specific moment.

25. Planned activity check (PLACHECK) is a variation of momentary time sampling in which the observer records whether each individual in a group is engaged in the target behavior.

Measuring Behavior by Permanent Products

26. Measuring behavior after it has occurred by measuring its effects on the environment is known as measurement by permanent product.

27. Measurement of many behaviors can be accomplished with contrived permanent products.

28. Measurement by permanent product offers numerous advantages: The practitioner is free to do other tasks; it enables the measurement of behaviors that occur at inconvenient or inaccessible times and places; measurement may be more accurate, complete, and continuous; it facilitates the collection of interobserver agreement and treatment integrity data; and it enables the measurement of complex behaviors and multiple response classes.

29. If moment-to-moment treatment decisions must be made during the session, measurement by permanent product may not be warranted.

30. Behaviors suitable for measurement with permanent products must meet two rules. Rule 1: Each occurrence of the target behavior must produce the same permanent product. Rule 2: The permanent product can be produced only by the target behavior.

Measurement Tools

31. Low tech tools are functional, easy to use, inexpensive, and effective for research, education, and treatment.

32. High-tech, digital and computer hardware and software systems for behavioral measurement and data analysis have become increasingly sophisticated and easier to use.

33. Developers have produced data collection and analysis software for observational measurement using laptops, handheld computers, personal digital assistants (PDAs), and desktop computers. These measurement tools provide the researcher and practitioner with advantages of practical value, efficiency, and ease of use.

Selecting a Measurement Method

34. When selecting a measurement method, consider the behavior change goals and expected direction of behavior change, the relative ease of detecting occurrences of the behavior, the environments where and times when the behavior will be measured, and the availability and skills of personnel who will observe and record the behavior.

KEY TERMS

artifact
celeration
count
discrete trial
duration
event recording
free operant
frequency
interresponse time (IRT)
latency
magnitude
measurement
measurement by permanent product
momentary time sampling
partial-interval recording
percentage
planned activity check (PLACHECK)
rate
repeatability
temporal extent
temporal locus
time sampling
topography
trials-to-criterion
whole-interval recording

MULTIPLE-CHOICE QUESTIONS

1. In order to be understood by everyone, measurement units must be:
   a. Analytic
   b. Subjective
   c. Labeled
   d. Unambiguous
   Hint: (See "Definition and Functions of Measurement in Applied Behavior Analysis")

2. Practitioners need measurement in order to do all of the following except
   a. Operationalize empiricism
   b. Identify current behavior levels
   c. Evaluate treatment
   d. Make accurate decisions
   Hint: (See "Definition and Functions of Measurement in Applied Behavior Analysis")

3. The dimensional quantity of duration is described as:
   a. Temporal locus
   b. Repeatability
   c. Temporal extent
   d. Count
   Hint: (See "Measurable Dimensions of Behavior")

4. _________ is obtained by combining observation time with a tally of the number of occurrences of behavior.
   a. Response latency
   b. Rate
   c. Celeration
   d. Whole interval
   Hint: (See "Measurable Dimensions of Behavior")

5. All of the following are procedures for measuring behavior except
   a. Event recording
   b. Timing
   c. Time sampling
   d. Observation
   Hint: (See "Measurement Tools")

6. Devices such as wrist counters, masking tape, buttons, and calculators are all useful when using ________ recording procedures.
   a. Event
   b. Timing
   c. Rate
   d. Duration
   Hint: (See "Measurement Tools")

7. The procedure that measures behavior after it has occurred by measuring its effect on the environment is called:
   a. Measurement artifact
   b. Planned activity check
   c. Permanent product
   d. Time sampling
   Hint: (See "Measuring Behavior by Permanent Products")

8. Which of the following is a reason not to use permanent product measurement?
   a. Moment-to-moment decisions must be made
   b. The behavior has a direct effect on the environment
   c. The product can only be produced by the target behavior
   d. Each behavior occurrence results in the same product
   Hint: (See "Measuring Behavior by Permanent Products")

9. One advantage of using computer-based observation and measurement systems is
   a. Many practitioners are familiar with it
   b. Data can be clustered and analyzed easily
   c. It is less expensive
   d. It is less intrusive than traditional methods
   Hint: (See "Measurement Tools")

10. Computer-assisted measurement of behavior allows the observer to do all of the following except:
    a. Collect permanent products
    b. Record multiple behaviors simultaneously
    c. Aggregate data easily
    d. Analyze data from different perspectives
    Hint: (See "Measurement Tools")

11. You are preparing to observe Jennifer, a second-grade student who has difficulty following teacher direction and frequently wanders around the classroom. The target behavior for Jennifer is being out of her seat. What is the most appropriate method for measuring this target behavior?
    a. Event recording
    b. Interresponse time
    c. Duration
    d. Planned activity check
    Hint: (See "Measurement Tools")

12. Marco is learning to assemble computer components as part of his new job at a factory. Speed and accuracy are critical skills for him to learn in assembling his component. What is the most appropriate method for measuring his behavior?
    a. Whole interval recording
    b. Duration
    c. Permanent product
    d. Response latency
    Hint: (See "Measurement Tools")

ESSAY-TYPE QUESTIONS

1. Describe one of the functions of measurement for practitioners of applied behavior analysis.

   Hint: (See "Definition and Functions of Measurement in Applied Behavior Analysis")

2. Define the three fundamental dimensional qualities of behavior.

   Hint: (See "Measurable Dimensions of Behavior")

3. Describe a one-time sampling procedure for measuring behavior and the appropriate conditions for using that procedure.

   Hint: (See "Measurement Tools")

4. Describe three advantages of measuring behavior by permanent product.

   Hint: (See "Measuring Behavior by Permanent Products")

5. Give two reasons for the increasing popularity of computer-assisted measurement systems.

   Hint: (See "Measurement Tools")

6. John, a fourth-grade teacher, wants to collect data on task completion for seven of his students. John has 34 students and does not have access to an aide or teaching assistant. Identify the most appropriate recording procedure for John to use and explain why it is the best choice.

   Hint: (See "Measurement Tools")

NOTES

1. Measurement is necessary but not sufficient for scientific understanding. (See Chapter 7.)
2. Noting that **frequency** has been used in the behavior analysis literature to refer to both count and rate, a practice that may cause confusion and misunderstanding, Merbitz, Merbitz, and Pennypacker (2016) made a case for limiting the use of frequency to mean rate. Carr, Nosik, and Luke (2018) surveyed the usage of *frequency* in contemporary behavior-analytic journals and textbooks and found that "the predominant usage of *frequency* was as *count*, not *rate*" (p. 436). Carr et al. encouraged behavior analysts to use *frequency* as a synonym for *count*.
3. *Latency* is most often used to describe the time between the onset of an antecedent stimulus change and the initiation of a response. However, the term can be used to refer to any measure of the temporal locus of a response with respect to any type of antecedent event. See Johnston and Pennypacker (2009).
4. Because percentages are ratios based on the same dimensional quantity, the dimensional quantity is canceled out and no longer exists in the percentage. For example, an accuracy percentage created by dividing the number of correct responses by the number of response opportunities removes the actual count. However, ratios created from different dimensional quantities retain the dimensional quantities of each component. For example, rate retains a count per unit of time. See Johnston and Pennypacker (2009) for further explication.
5. When a person reports a percentage in excess of 100% (e.g., "Our mutual fund grew by 120% during the recent bear market"), he is probably using the misstated percentage in *comparison to* the previous base unit, not as a proportion of it. In this example, the mutual fund's 20% increase makes its value 1.2 times greater than its value when the bear market began.
6. Researchers sometimes manipulate and control response topography and magnitude to assess their possible effects as independent variables. Piazza, Roane, Kenney, Boney, and Abt (2002) analyzed the effects of different response topographies on the occurrence of pica (e.g., ingesting nonnutritive matter that may threaten one's life) by three females. Pica items were located in various places, which required the subjects to respond in different ways (e.g., reach, bend over, get on the floor, open a container) to obtain them. Pica decreased when more elaborate response topographies were required to obtain pica items. Van Houten (1993) reported that when a boy with a long history of intense and high occurrences of face slapping wore 1.5-pound wrist weights, face slapping immediately dropped to zero. Studies such as these suggest that problem behaviors may be reduced when engaging in those behaviors requires more effortful responses in terms of topography or magnitude (Friman & Poling, 1995).
7. A variety of terms appear in the applied behavior analysis literature to describe measurement procedures that involve observing and recording behavior within or at the end of specified intervals. Some authors use *time sampling* to refer only to momentary time sampling. We consider whole-interval and partial-interval recording as time sampling methods because the data they produce provide a representative "sample" of the target behavior during the observation period.
8. Edwards and Christophersen (1993) described a time-lapse videotape recorder (TLVCR) that automatically records on one 2-hour tape time samples of behavior over observation periods ranging from 2 to 400 hours. A TLVCR programmed to record over a 12-hour period would record for 0.10 of each second. Such a system can be useful for recoding very low rate behaviors and for behaviors that occur over long periods (e.g., a child's sleep behavior).
9. Descriptions of a variety of computer-assisted behavioral measurement systems can be found in Bullock, Fisher, and Hagopian (2001); Emerson, Reever, and Felce (2000); Gravlee, Zenk, Woods, Rowe, and Schultz (2006); Kahng & Iwata (1998, 2000); Noldus, Trienes, Hendriksen, and Jansen (2000); Sidenar, Shabani, and Carr (2004); Tapp & Walden (2000); and Tapp and Wehby (2000).

CHAPTER 5

Improving and Assessing the Quality of Behavioral Measurement

LEARNING OBJECTIVES

- State and describe the elements of useful scientific measurement.
- Identify threats to measurement validity.
- Describe various threats to the accuracy and reliability of measurement.
- Identify and explain ways to assess the accuracy and reliability of behavioral measurement.
- Identify and explain how to assess interobserver agreement (IOA) for a variety of data sets.

Three friends—John, Tim, and Bill—took a bicycle ride together. At the end of the ride John looked at his handlebar-mounted bike computer and said, "We rode 68 miles. Excellent!" "My computer shows 67.5 miles. Good ride, fellas!" Tim replied. As he dismounted and rubbed his rump, the third biker, Bill, said, "Gee whiz, I'm sore! We must've ridden 100 miles!" A few days later, the three friends completed the same route. After the second ride, John's computer showed 68 miles, Tim's computer read 70 miles, and Bill, because he wasn't quite as sore as he was after the first ride, said they had ridden 90 miles. Following a third ride on the same country roads, John, Tim, and Bill reported distances of 68, 65, and 80 miles, respectively.

The data obtained by measuring behavior are the primary material with which behavioral researchers and practitioners guide and evaluate their work. Applied behavior analysts measure socially significant behaviors to help determine which behaviors need to be changed, to detect and compare the effects of various interventions on behaviors targeted for change, and to evaluate the acquisition, maintenance, and generalization of behavior changes.

Because so much of what the behavior analyst does either as a researcher or as a practitioner depends on measurement, concerns about the legitimacy of the data it produces must be paramount. Do the data meaningfully reflect the original reason(s) for measuring the behavior? Do the data represent the true extent of the behavior as it actually occurred? Do the data provide a consistent picture of the behavior? In other words, can the data be trusted?

Chapter 4 identified the measurable dimensions of behavior and described the measurement methods most often used in applied behavior analysis. This chapter focuses on improving and assessing the quality of behavioral measurement. We begin by defining the essential indicators of trustworthy measurement: validity, accuracy, and reliability. Next, common threats to measurement are identified and suggestions for combating these threats are presented. The chapter's final sections detail procedures for assessing the accuracy, reliability, and believability of behavioral measurement.

INDICATORS OF TRUSTWORTHY MEASUREMENT

How trustworthy were the measures reported by the three bicyclists in the opening anecdote? Which of the three friends' data would contribute to a scientific account of the miles they had ridden? To contribute to a scientific account of events, measurement must be valid, accurate, and reliable. Were the three friends' measurements characterized by validity, accuracy, and reliability?

Validity

Measurement has **validity** when it yields data directly relevant to the phenomenon measured and to the reason(s) for measuring it. Determining the validity of measurement revolves around this basic question: Was a relevant dimension of the behavior that is the focus of the investigation measured directly and legitimately?

Did the measurements of miles ridden by the three bicyclists have validity? Because the bikers wanted to know how far they had ridden each time, number of miles ridden was a relevant, or valid, dimension of their riding behavior. Had the bikers' primary interest been how long or how fast they had ridden, the number of miles ridden would not have been a valid measure. John and Tim's use of their bike computers to measure directly the miles they rode was a valid measure. Because Bill used an indirect measure (the relative tenderness of his backside)

to determine the number of miles he had ridden, the validity of Bill's mileage data is suspect. A direct measure of the actual behavior of interest will always possess more validity than an indirect measure, because a direct measure does not require an inference about its relation to the behavior of interest, whereas an indirect measure always requires such an inference. Although soreness may be related to the distance ridden, because it is also influenced by such factors as the time on the bike saddle, the roughness of the road, riding speed, and how much (or little) a person has ridden recently, soreness as a measure of mileage has little validity.

Valid measurement in applied behavior analysis requires three equally important elements: (a) measuring directly a socially significant target behavior (see Chapter 3), (b) measuring a dimension (e.g., rate, duration) of the target behavior relevant to the question or concern about the behavior (see Chapter 4), and (c) ensuring that the data are representative of the behavior's occurrence under conditions and during times that are most relevant to the question or concern about the behavior. When any of these elements are suspect or lacking—no matter how accurate and reliable was the measurement that produced the data—the validity of the resultant data is compromised, perhaps to the point of being meaninglessness.

Accuracy

When used in the context of measurement, **accuracy** refers to the extent to which the **observed value** (i.e., the quantitative labels produced by measuring an event) match the true value of the event. In other words, measurement is accurate to the degree that it corresponds to the true value of the thing measured. Obtaining a **true value** requires special observation and recording procedures that "must be at least somewhat different from those used to produce the data being evaluated, and the differences must have the effect of minimizing the possibility of error to an uncommon degree" (Johnston & Pennypacker, 2009, pp. 136–137).

How accurate were the three bikers' measures of miles ridden? Because each biker reported a different measure of the same event, all of their data could not be accurate. Skeptical of the training miles the three cyclists were claiming, a friend of theirs, Lee, drove the same country roads with GPS-calibrated odomenter in his car. At the end of the route, the odometer read 58 miles. Using the measure obtained by the odometer as the true value of the route's distance, Lee determined that none of the three cyclists' measures was accurate. Each rider had overestimated the true mileage.

By comparing the mileage reported by John, Tim, and Bill with the true value of the route's distance, Lee discovered not only that the riders' data were inaccurate but also that the data reported by all three riders were contaminated by a particular type of measurement error called measurement bias. **Measurement bias** refers to nonrandom measurement error—that an error in measurement that is likely to be in one direction. Random measurement error is just as likely to overestimate the true value of an event as it is to underestimate it. Because John, Tim, and Bill consistently overestimated the actual miles they had ridden, their data contained measurement bias.

Reliability

Reliability refers to the consistency of measurement, specifically, the extent to which repeated measurement of the same event yields the same value. In other words, reliable measurement is consistent measurement. Like validity and accuracy, reliability is a relative concept; it is a matter of degree. The closer the values obtained by repeated measurement of the same event are to one another, the greater the reliability. Conversely, the more observed values from repeated measurement of the same event differ from one another, the less the reliability.

How reliable were the bicyclists' measurements? Because John obtained the same value, 68 miles, each time he measured the same route, his measurement had complete reliability. Tim's three measures of the same ride—67.5, 70, and 65 miles—differed from one another by as much as 5 miles. Therefore, Tim's measurement was less reliable than John's. Bill's measurement system was the least reliable of all, yielding values for the same route ranging from 80 to 100 miles.

Relative Importance of Validity, Accuracy, and Reliability

Behavioral measurement should provide legitimate data for evaluating behavior change and guiding research and treatment decisions. Data of the highest quality (i.e., data that are most useful and trustworthy for advancing scientific knowledge or for guiding practice) are produced by valid, accurate, and reliable measurement (see Figure 5.1). Validity, accuracy, and reliability are relative concepts; each can range from high to low.

Measurement must be both valid and accurate for the data to be trustworthy. If measurement is not valid, accuracy is moot. Accurately measuring a behavior that is not the focus of the investigation, accurately measuring an irrelevant dimension of the target behavior, or accurately measuring the behavior under circumstances or at times not representative of the conditions and times relevant to the analysis will yield invalid data. Conversely, the data obtained from measuring a meaningful dimension of the right behavior under the relevant circumstances and times are of little use if the observed values provide an inaccurate picture of the behavior. Inaccurate measurement renders invalid the data obtained by otherwise valid measurement.

Reliability should never be confused with accuracy. Although John's bicycle computer provided totally reliable measures, it was also totally inaccurate. Concern about the reliability of data in the absence of a prior interest in their accuracy suggests that reliability is being mistaken for accuracy. The question for a researcher or someone who is reading a published study is not "Are the data reliable?" but "Are the data accurate?" (Johnston & Pennypacker, 1993a, p. 146)

If accuracy trumps reliability—and it does—why should researchers and practitioners be concerned with the reliability of measurement? Although high reliability does not mean high accuracy, poor reliability reveals problems with accuracy. Because Tim and Bill's measurements were not reliable, we know that at least some of the data they reported could not be

Figure 5.1 Valid, accurate, and reliable measurement yields the most trustworthy and useful data for science and science-based practice.

Measurement that is . . .

Valid	Accurate	Reliable	. . . yields data that are . . .
Yes	Yes	Yes	. . . most useful for advancing scientific knowledge and guiding data-based practice.
No	Yes	Yes	. . . meaningless for the purposes for which measurement was conducted.
Yes	No	Yes	. . . always wrong.[a]
Yes	Yes	No[b]	. . . sometimes wrong.[c]

a. If adjusted for consistent measurement error of standard size and direction, inaccurate data may still be usable.

b. If the accuracy of every datum in a data set can be confirmed, reliability is a moot point. In practice, however, that is seldom possible; therefore, knowing the consistency with which observers applied a valid and accurate measurement system contributes to the level of confidence in the overall trustworthiness of the data set.

c. User is unable to separate good data from bad.

accurate, knowledge that could and should lead to checking the accuracy of their measurement tools and procedures.

Highly reliable measurement means that whatever degree of accuracy (or inaccuracy) exists in the measurement system will be revealed consistently in the data. If it can be determined that John's computer reliably obtains observed values higher than the true values by a constant amount or proportion, the data could be adjusted to accommodate for that constant degree of inaccuracy.

The next two sections of the chapter describe methods for combating common threats to the validity, accuracy, and reliability of behavioral measurement.

THREATS TO VALID MEASUREMENT

The validity of behavioral data is threatened when measurement is indirect, when the wrong dimension of the target behavior is measured, or when measurement is conducted in such a way that the data it produces are an artifact of the actual events.

Indirect Measurement

Indirect measurement occurs when the researcher or practitioner measures a proxy, or stand-in, for the actual behavior of interest. Indirect measurement provides secondhand or "filtered" information that requires the researcher or practitioner to make inferences about the relationship between the event that was measured and the actual behavior of interest (Komaki, 1998). For example, a student's responses to a questionnaire designed to measure his socialization with peers is an indirect measure. Using a student's score on a standardized math achievement test as an indicator of mastery is another example of indirect measurement. Both of these illustrations require an inference about the behavior being measured and the actual behavior of interest. By contrast, a student's score on a properly constructed test consisting of math problems from recently covered curriculum content is a direct measure requiring no inferences about what it means with respect to her performance in the curriculum (e.g., Hosp & Hosp, 2003; Stecker, Fuchs, & Fuchs, 2005). Conversely, **direct measurement** occurs when the behavior measured is exactly the same as the behavior that is the focus of the investigation or behavior change program.

Indirect measurement rarely occurs in applied behavior analysis because meeting the applied dimension of ABA includes the targeting and meaningful (i.e., valid) measurement of socially significant behavior. Sometimes, however, the researcher or practitioner has no direct and reliable access to the behavior of interest and so must use some form of indirect measurement. Because researchers studying adherence to medical regimens cannot directly observe and measure patients' behavior in their homes, they rely on self-reports for their data (e.g., La Greca & Schuman, 1995).[1]

Indirect measurement is sometimes used to make inferences about private events or affective states. For example, Green and Reid (1996) used direct measures of smiling to represent "happiness" by persons with profound multiple disabilities. However, research on private events does not necessarily involve indirect measurement. A research participant who has been trained to observe his own private events is measuring the behavior of interest directly (e.g., Kostewicz, Kubina, & Cooper, 2000; Kubina, Haertel, & Cooper, 1994).

Whenever indirect measurement is used, it is the researcher's responsibility to provide evidence that the event measured directly reflects, in some reliable and meaningful way, something about the behavior for which the researcher wishes to draw conclusions (Johnston & Pennypacker, 2009). In other words, it is incumbent upon the researcher to provide a convincing case for the validity of her data. Although it is sometimes attempted, the case for validity cannot be achieved by simply attaching the name of the thing one claims to be measuring to the thing actually measured. With respect to that point, Marr (2003) recounted this anecdote about Abraham Lincoln:

> "Sir, how many legs does this donkey have?"
> "Four, Mr. Lincoln."
> "And how many tails does it have?"
> "One, Mr. Lincoln."
> "Now, sir, what if we were to call a tail a leg; how many legs would the donkey have?"
> "Five, Mr. Lincoln."
> "No sir, for you cannot make a tail into a leg by calling it one." (pp. 66–67)

Measuring the Wrong Dimension of the Target Behavior

The validity of behavioral measurement is threatened much more often by measuring the wrong dimension of the behavior of interest than it is by indirect measurement. Valid measurement yields data that are relevant to the questions about the behavior one seeks to answer through measurement. Validity is compromised when measurement produces values for a dimension of the behavior ill suited for, or irrelevant to, the reason for measuring the behavior.

Johnston and Pennypacker (1980) provided an excellent example of the importance of measuring a dimension that fits the reasons for measurement. "Sticking a ruler in a pot of water as the temperature is raised will yield highly reliable measures of the depth of the water but will tell us very little about the changing temperature" (p. 192). While the units of measurement on a ruler are well suited for measuring length, or in this case, depth, they are not at all valid for measuring temperature. If the purpose of measuring the water is to determine whether it has reached the ideal temperature for making a pot of tea, a thermometer is the correct measurement tool.

If you are interested in measuring a student's academic endurance with oral reading, counting the number of correct and incorrect words read per minute without measuring and reporting the total time that the student read will not provide valid data on endurance. Number of words read per minute alone does not fit the reason for measuring reading (i.e., academic endurance). To measure endurance, the practitioner would need to report the duration of the reading period (e.g., 30 minutes). Similarly, measuring the percentage of trials on which a student makes a correct response would not provide valid data for answering questions about the student's developing fluency with a skill, whereas measuring the number of correct responses per minute and the changing rates of responding (celeration) would.

Measurement Artifacts

Directly measuring a relevant dimension of a socially significant target behavior does not guarantee valid measurement. Validity is reduced when the data—no matter how accurate or reliable they are—do not give a meaningful (i.e., valid) representation of the behavior. When data give an unwarranted or misleading picture of the behavior because of the way measurement was conducted, the data are called an *artifact*. As introduced in Chapter 4, a *measurement artifact* is something that appears to exist because of the way it is measured. Discontinuous measurement, poorly scheduled measurement periods, and using insensitive or limiting measurement scales are common causes of measurement artifacts.

Discontinuous Measurement

Because behavior is a dynamic and continuous phenomenon that occurs and changes over time, continuous measurement is the gold standard in behavior analysis research and practice. **Continuous measurement** is measurement that detects all instances of the target behavior during the observation period. The development of automated data collection technologies has produced an increasing trend of continuous measurement in behavioral research (Kostewicz, King, Datchuk, & Brennan, 2016).

Discontinuous measurement is any form of measurement in which some instances of the response class(es) of interest may not be detected. A data set produced by discontinuous measurement—no matter how accurate and reliable—may be an artifact of the measurement system rather than accurate portrayal of behavioral events.

In the absence of automatic measurement technology, continuous measurement may be prohibited by factors such as the nature of the target behavior being measured; the number of participants or clients whose behavior must be measured; the number, duration, and scheduling of observation periods; and the availability of trained personnel to observe and collect the data. In such instances, applied behavior analysts typically use one of the three primary methods of discontinuous measurement described in Chapter 4 (partial-interval recording, whole-interval recording, or momentary time sampling).

A study by Thomson, Holmber, and Baer (1974) demonstrated the extent of artificial variability in a data set that may be caused by discontinuous measurement. A single, highly experienced observer used three different procedures for scheduling time sampling observations to measure the behavior of four subjects (two teachers and two children) in a preschool setting during 64-minute sessions. Thomson and colleagues called the three time sampling procedures *contiguous*, *alternating*, and *sequential*. With each time sampling procedure, one fourth of the observer's time (i.e., 16 minutes) was assigned to each of the four subjects.

When the contiguous observation schedule was used, the observer recorded the behavior of Subject 1 throughout the first 16 minutes of the session, recorded the behavior of Subject 2 during the second 16 minutes, and so on until all four students had been observed. In the alternating mode, Subjects 1 and 2 were observed in alternating intervals during the first half of the session, and Subjects 3 and 4 were observed in the same fashion during the last half of the session. Specifically, Subject 1 was observed during the first 4 minutes, Subject 2 during the next 4 minutes, Subject 1 during the next 4 minutes, and so on until 32 minutes had expired. The same procedure was then used for Subjects 3 and 4 during the last 32 minutes of the session. The sequential approach systematically rotated the four subjects through 4-minute observations. Subject 1 was observed during the first 4 minutes, Subject 2 during the second 4 minutes, Subject 3 during the third 4 minutes, and Subject 4 during the fourth 4 minutes. This sequence was repeated 4 times to give the total of 64 minutes of observation.

To arrive at the percentage of artifactual variance in the data associated with each time sampling schedule, Thomson and colleagues (1974) compared the observer's data with "actual rates" for each subject produced by continuous measurement of each subject for the same 64-minute sessions. Results of the study showed clearly that the contiguous and alternating schedules produced the most unrepresentative (and, therefore, less valid) measures of the target behaviors (often more than 50% variance from continuous measurement), whereas sequential sampling procedure produced results that more closely resembled the data obtained through continuous recording (from 4% to 11% variance from continuous measurement) (cf. Tiger et al., 2013).

In spite of its inherent limitations, discontinuous measurement is used in many studies in applied behavior analysis in which individual observers measure the behavior of multiple subjects within the same session. Minimizing the threat to validity posed by discontinuous measurement requires careful consideration of when observation and measurement periods should be scheduled. Infrequent measurement, no matter how accurate and reliable it is, often yields results that are an artifact. Although a single measure reveals the presence or absence of the target behavior at a given point in time, it may not be representative of the typical value for the behavior.[2] As a general rule, observations should be scheduled on a daily or frequent basis, even if for only brief periods.

Ideally, all occurrences of the behavior of interest should be recorded. However, when available resources preclude continuous measurement throughout an observation period, the use of sampling procedures is necessary. A sampling procedure may be sufficient for decision making and analysis if the samples represent a valid approximation of the true parameters of the behavior of interest. When measurement cannot be continuous throughout an observation period, it is generally preferable to sample the occurrence of the target behavior for numerous brief observation intervals that are evenly distributed throughout the session than it is to use longer, less frequent intervals (Thomson et al., 1974; Thompson, Symons, & Felce, 2000). For example, measuring a subject's behavior in thirty 10-second intervals equally distributed within a 30-minute session will likely yield more representative data than will observing the person for a single 5-minute period during the half hour.

Measuring behavior with observation intervals that are too short or too long may result in data that grossly over- or underestimates the true occurrence of behavior. For example, measuring off-task behavior by partial-interval recording with 10-minute intervals may produce data that make even the most diligent of students appear to be highly off task.

Poorly Scheduled Measurement Periods

The observation schedule should be standardized to provide an equal opportunity for the occurrence or nonoccurrence of the behavior across sessions and consistent environmental conditions from one observation session to the next. When neither of these requirements is met, the resultant data may not represent actual events and therefore be invalid. If observation periods are scheduled at times when and/or places where the frequency of behavior is atypical, the data may not represent periods of high or low responding. For example, measuring students' being on task during only the first 5 minutes of each day's 20-minute cooperative learning group activity may yield data that make on-task behavior appear higher than it actually is over the entire activity.

When data will be used to assess the effects of an intervention or treatment, the most conservative observation times should be selected. That is, target behaviors should be measured during those times when their rates of occurrence is most likely to be different from the desired or predicted outcomes of the treatment. Measurement of behaviors targeted for reduction should occur during times when those behaviors are most likely to occur at their highest response rates. Conversely, behaviors targeted for increase should be measured when high-rate responding is least likely. If an intervention is not planned—as might be the case in a descriptive study—it is important to select the observation times most likely to yield data that are generally representative of the behavior.

Insensitive and/or Limited Measurement Scales

Artifactual data may result from using measurement scales that cannot detect the full range of relevant values or that are insensitive to meaningful changes in behavior. Data obtained with a measurement scale that does not detect the full range of relevant performances may incorrectly imply that behavior cannot occur at levels below or above obtained measures because the scale has imposed an artificial floor or ceiling on performance. For example, measuring a student's oral reading fluency by giving him a 100-word passage to read in 1 minute may yield data that suggest that his maximum performance is 100 wpm.

Measurement scales that are over- or under-sensitive to relevant changes in behavior may produce misleading data showing that meaningful behavior change has (or has not) occurred. For example, using a percentage measure scaled in 10% increments to evaluate the effects of an intervention to improve quality control in a manufacturing plant may not reveal important changes in performance if improvement in the percentage of correctly fabricated widgets from a baseline level of 92% to a range of 97% to 98% is the difference between unacceptable and acceptable (i.e., profitable) performance.

THREATS TO ACCURATE AND RELIABLE MEASUREMENT

Human error is the biggest threat to the accuracy and reliability of data in applied behavior analysis. Unlike the experimental analysis of behavior, in which measurement is typically automated, most investigations in applied behavior analysis use human observers to measure behavior.[3] Factors that contribute to human measurement error include poorly designed measurement systems, inadequate observer training, and expectations about what the data should reveal.

Poorly Designed Measurement System

Unnecessarily cumbersome and difficult-to-use measurement systems create needless loss of accuracy and reliability. Collecting behavioral data in applied settings requires attention, keen judgment, and perseverance. The more taxing and difficult a measurement system is to use, the less likely an observer will be to consistently detect and record all instances of the target behavior. Simplifying the measurement system as much as possible minimizes measurement errors.

The complexity of measurement includes such variables as the number of individuals observed, the number of behaviors recorded, the duration of observation periods, and/or the duration of the observation intervals, all of which may affect the quality of measurement. For instance, observing several individuals is more complex than observing one person; recording several behaviors is more complex than recording a single behavior; and using contiguous 5-second observation intervals with no time between intervals to record the results of the observation is more difficult than a system in which time is reserved for recording data.

Specific recommendations for reducing complexity depend on the particular nature of the study. However, when using time sampling measurements, applied behavior analysts can consider modifications such as decreasing the number of simultaneously observed individuals or behaviors, decreasing the duration of the observation sessions (e.g., from 30 minutes to 15 minutes), and increasing the duration of time intervals (e.g., from 5 to 10 seconds). Requiring more practice during observer training, establishing a higher criterion for mastery of the observational code, and providing more frequent feedback to observers may also reduce the possible negative effects of complex measurement.

Inadequate Observer Training

Careful attention must be paid to the selection and training of observers. Explicit and systematic observer training is essential for the collection of trustworthy data. Observation and coding systems require observers to discriminate the occurrence and nonoccurrence of specific classes of behaviors or events against an often complex and dynamic background of other behaviors or events and to record their observations onto a data sheet. Observers must learn the definitions for each behavior or event to be measured; a coded notation system or icon for each event to be measured; a common set of recording procedures (e.g., marking a data collection form, keystrokes, touching icons); and how to correct recording errors (e.g., writing a plus sign instead of a minus sign, hitting the F6 key instead of the F5 key, touching the 'on-task' icon instead of the 'off-task' icon).

Selecting Observers Carefully

Admittedly, applied researchers often scramble to find data collectors, but not all volunteers should be accepted into training. Potential observers should be interviewed to determine past experiences with observation and measurement activities, current schedule and upcoming commitments, work ethic and motivation, and overall social skills. The interview might include a pretest to determine current observation and skill levels. This can be accomplished by having potential observers watch short video clips of behaviors similar to what they may be asked to observe and noting their performance against a criterion.

Training Observers to an Objective Standard of Competency

Observer trainees should meet a specified criterion for recording before conducting observations in applied settings. During training, observers should practice recording numerous examples and nonexamples of the target behavior(s) and receive a critique and performance feedback. Observers should have numerous practice sessions before actual data collection. Training should continue until a predetermined criterion is achieved (e.g., 95% accuracy for two or three consecutive sessions). For example, in training observers to measure the completion of preventive maintenance tasks of heavy equipment by military personnel, Komaki (1998) required three consecutive sessions of at least 90% agreement with a true value.

Various methods can be used to train observers. These include sample vignettes, narrative descriptions, video modeling sequences (see Chapter 21), role playing, and practice sessions in the environment in which actual data will be collected. Practice sessions in natural settings are especially beneficial because they allow both observers and participants to adapt to each other's presence and may reduce the reactive effects of the presence of observers on participants' behavior. For example, observer trainees for a study that measured appropriate and inappropriate recruiting responses by middle school students with learning disabilities during collaborative learning group (CLG) activities and three types of peer reactions to those responses (instructional feedback, praise, and negative statements) were required to first demonstrate accurate recording (using experimenters' descriptions of events as true values) and then practice observing and measuring those events in a classroom similar to the experimental setting:

> The first author created two observation training scripts, each script consisting of 10 vignettes of various CLG interactions the observers were likely to encounter. The vignettes included appropriate and inappropriate recruiting responses and various types of peer reactions to recruiting. The first author discussed the first script with the observers, demonstrating how each vignette should be marked on the recording form. Each observer then independently completed a recording form based on the second script. When the observers achieved at least 95% interobserver agreement (IOA) with the first author on the frequency and type of student recruiting and peer assistance, they began practicing data collection in the language arts classroom during CLG activities. Baseline data collection began when the observers attained 95% IOA with the first author in the classroom. (Wolford, Alber, & Heward, 2001, p. 164)

The following steps are an example of a systematic approach for training observers.

Step 1 Trainees read the target behavior definitions and become familiar with data collection forms, procedures for recording their observations, and the proper use of any measurement or recording devices (e.g., tape recorders, stopwatches, laptops, PDAs, bar code scanners).

Step 2 Trainees practice recording simplified narrative descriptions of behavioral vignettes until they obtain 100% accuracy over a predetermined number of instances.

Step 3 Trainees practice recording longer, more complex narrative descriptions of behavioral vignettes until they obtain 100% accuracy for a predetermined number of episodes.

Step 4 Trainees practice observing and recording data from videorecorded or role-played vignettes depicting the target behavior(s) at the same speed and complexity as will occur in the natural environment. Training vignettes should be scripted and sequenced to provide trainees practice in making increasingly difficult discriminations between the occurrence and nonoccurrence of the target behavior(s). Having trainees rescore the same series of vignettes a second time and comparing the reliability of their measures provides an assessment of the consistency with which the trainees are applying the measurement system. Trainees remain at this step until their data reach pre-established accuracy and reliability criteria. (If the study involved collecting data from natural permanent products such as compositions or academic worksheets, Steps 2 through 4 should provide trainees with practice scoring increasingly extensive and more difficult-to-score examples.)

Step 5 The final training step of observer training is practice in collecting data in the natural environment. An experienced observer accompanies the trainee and simultaneously and independently measures the target behaviors. Each practice session ends with the trainee and experienced observer comparing their data sheets and discussing any questionable or heretofore unforeseen instances. Training continues until a pre-established criterion of agreement between the experienced observer and the trainee is achieved (e.g., at least 90% for three consecutive sessions).

Providing Ongoing Training to Minimize Observer Drift

Over the course of a study, observers sometimes alter, often unknowingly, the way they apply a measurement system. Called **observer drift**, these unintended changes in the way data are collected may produce measurement error. Observer drift usually entails a shift in the observer's interpretation of the definition of the target behavior from that used in training. Observer drift occurs when observers expand or compress the original definition of the target behavior. For example, observer drift might be responsible for the same behaviors by a child that were recorded by an observer as instances of noncompliance during the first week of a study being scored as instances of compliance during the study's final week. Observers are usually unaware of the drift in their measurement.

Occasional observer retraining or booster sessions throughout the investigation can minimize observer drift. Continued training provides the opportunity for observers to receive frequent feedback on the accuracy and reliability of measurement. Ongoing training can occur at regular, prescheduled intervals (e.g., every Friday morning) or randomly.

Unintended Influences on Observers

Ideally, data reported by observers have been influenced only by the actual occurrences and nonoccurrences of the target behavior(s) observers have been trained to measure. In reality, however, a variety of unintended and undesired influences on observers can threaten the accuracy and reliability of the data they report. Common causes of this type of measurement error include an observer's presuppositions about the expected outcomes of the data and awareness that others are measuring the same behavior.

Observer Expectations

Observer expectations that the target behavior should occur at a certain level under particular conditions, or change when a change in the environment has been made, pose a major threat to accurate measurement. For example, if an observer believes or predicts that a teacher's implementation of a token economy should decrease inappropriate student behavior, she may record fewer inappropriate behaviors during the token reinforcement condition than she would have recorded otherwise without holding that expectation. Data influenced by an observer's expectations or efforts to obtain results that will please the researcher are characterized by measurement bias.

The surest way to minimize measurement bias caused by observer expectations is to use naive observers. A totally **naive observer** is a trained observer who is unaware of the study's purpose and/or the experimental conditions in effect during a given phase or observation period. Researchers should inform observer trainees that they will receive limited information about the study's purpose and why that is. However, maintaining observers' naiveté is often difficult and sometimes impossible.

When observers are aware of the purpose or hypothesized results of an investigation, measurement bias can be minimized by using target behavior definitions and recording procedures that will give a conservative picture of the behavior (e.g., whole-interval recording of on-task behavior with 10-second rather than 5-second intervals), frank and repeated discussion with observers about the importance of collecting accurate data, and frequent feedback to observers on the extent to which their data agree with true values or data obtained by naive observers. Observers should not receive feedback about the extent to which their data confirm or run counter to hypothesized results or treatment goals.

Observer Reactivity

Measurement error resulting from an observer's awareness that others are evaluating the data he reports is called **observer reactivity**. Like reactivity that may occur when

participants are aware that their behavior is being observed, the behavior of observers (i.e., the data they record and report) can be influenced by the knowledge that others are evaluating the data. For example, knowing that the researcher or another observer is watching the same behavior at the same time, or will monitor the measurement through video or audio recording later, may produce observer reactivity. If the observer anticipates that another observer will record the behavior in a certain way, his data may be influenced by what he anticipates the other observer may record.

Monitoring observers as unobtrusively as possible on an unpredictable schedule helps reduce observer reactivity. Separating multiple observers by distance or partition reduces the likelihood that their measures will be influenced by one another's during an observation. One-way mirrors in some research and clinical settings eliminate visual contact between the primary and secondary observers. If audio or video recordings are made of sessions, the secondary observer can measure the behavior at a later time and the primary observer never has to come into contact with the secondary observer. In settings where one-way mirrors are not possible, and where audio or video recording may be intrusive, the secondary observer might begin measuring the behavior at a time unknown to the primary observer. For example, if the primary observer begins measuring behavior with the first interval, the secondary observer could start measuring behavior after 10 minutes have elapsed. The intervals used for comparisons would begin at the 10-minute mark, ignoring those intervals that the primary observer recorded beforehand.

ASSESSING THE ACCURACY AND RELIABILITY OF BEHAVIORAL MEASUREMENT

After designing a measurement system that will produce a valid representation of the target behavior and training observers to use it in a manner likely to yield accurate and reliable data, the researcher's next measurement-related tasks are evaluating the extent to which the data are, in fact, accurate and reliable. Essentially, all procedures for assessing the accuracy and reliability of behavioral data entail some form of "measuring the measurement system."

Assessing the Accuracy of Measurement

Measurement is accurate when the observed values (i.e., the numbers obtained by measuring an event) match the true values of the event. The fundamental reason for determining the accuracy of data is obvious: No one wants to base research conclusions or make treatment decisions on faulty data. More specifically, conducting accuracy assessments serves four interrelated purposes. First, it is important to determine early in an analysis whether the data are good enough to serve as the basis for making experimental or treatment decisions. The first person the researcher or practitioner must try to convince that the data are accurate is herself. Second, accuracy assessments enable the discovery and correction of specific instances of measurement error. The two other approaches to assessing the quality of data, to be discussed later in this chapter—reliability assessments and interobserver agreement—can alert the behavior analyst to the likelihood of measurement errors, but neither approach identifies errors. Only the direct assessment of measurement accuracy allows practitioners or applied researchers to detect and correct faulty data.

A third reason for conducting accuracy assessments is to reveal consistent patterns of measurement error, which can lead to calibration of the measurement system. **Calibration** entails comparing the data produced by a measurement system to a known standard or true value and, if necessary, adjusting the measurement system so the data it produces match the known standard. For example, knowing that John's bicycle computer reliably obtained a measure of 68 miles for a route with a true value of 58 miles led not only to the cyclists correcting the data in hand (in this case, confessing to one another and to their friend Lee that they had not ridden as many miles as previously claimed) but to their calibrating John's bicycle computer (in this case, by adjusting the wheel circumference setting) so that future measures would be accurate.

Mudford, Zeleny, Fisher, Klum, and Owen (2011) calibrated human observers by comparing their data to criterion (known) values for the same events. Five experienced and five novice observers participated in the study. Observers viewed scripted video records of two adults playing the roles of therapist and client and recorded three behaviors (client pinches therapist, model prompts by the therapist, and verbal+gestural prompts by the therapist) with laptop computers. Each observer recorded data from 10 different 10-minute samples viewed in random order. Each sample included a varied number of pinches (from 0 to 80) and 15 randomly distributed therapist prompts.

Because the video samples were scripted, the number of occurrences for each of the three target behaviors was a known (true) value. What the researchers called the "criterion records" for when each response occurred during the video samples to 1-second precision were obtained by three of the study's co-authors independently, noting times of occurrence while playing, pausing, rewinding, and replaying the video recordings. "The only discrepancies between independent records were on the exact second in which responses occurred. Discrepancies were resolved by authors reviewing the sample together until agreement on time of occurrence was achieved" (p. 575).

After using three different IOA methods and linear regression to compare each observer's data with "criterion records," Mudford et al. (2011) concluded that, "Experienced observers demonstrated almost uniformly high levels of agreement with criterion records, accuracy, and precision. . . . Novice observers, as a group, did not fall far behind the experienced observers in accuracy and precision" (p. 582). Although these results apply only to the measurement system and observers assessed, the study demonstrates a calibration method that may improve the accuracy of data collected by applied behavior analysts.

Calibrating any measurement tool, whether it is a mechanical device or human observer, entails comparing the data obtained by the tool against a true value. The measure obtained by the GPS-odometer served as the true value for calibrating John's bike computer. Calibration of a timing device such as a stopwatch or countdown timer could be made against a known

standard: the "ytterbium clock."[4] If no differences are detected when comparing the timing device against the ytterbium clock, or if the differences are tolerable for the intended purposes of measurement, then calibration is satisfied. If significant differences are found, the timing device would need to be reset to the standard. We recommend frequent accuracy assessments in the beginning stages of an analysis. Then, if the assessments have produced high accuracy, less frequent assessments can be conducted to check the calibration of the recorders.

A fourth reason for conducting accuracy assessments is to assure consumers that the data are accurate. Including the results of accuracy assessments in research reports helps readers judge the trustworthiness of the data being offered for interpretation.

Establishing True Values

True values for some behaviors are evident and universally accepted. For example, obtaining the true values of correct responses in academic areas such as math and spelling is straightforward. The correct response to the arithmetic problem 2 + 2 = ? has a true value of 4, and the *Oxford English Dictionary* is a source of true values for assessing the accuracy of measuring the spelling of English words.[5]

Although not universal, true values for many socially significant behaviors of interest to applied researchers and practitioners can be established conditionally on local context. For example, the correct response to the question "Name the three starches recommended as thickeners for pan gravy" on a quiz given to students in a culinary school has no universal true value. Nevertheless, a true value relevant to the students taking the quiz can be found in the instructor's course materials.

> There is no single way to obtain true values. There are only the requirements that: (a) the procedure must be different from that used to collect the data that are being evaluated; and (b) the procedure must incorporate extraordinary steps that avoid or remove possible sources of error. (Johnston & Pennypacker, 2009, p. 146)

True values for each of the preceding examples were obtained through sources independent of the measures to be evaluated. Establishing true values for many behaviors studied by applied behavior analysts is difficult because the process for determining a true value must be different from the measurement procedures used to obtain the data one wishes to compare to the true value.

This is difficult, true, but not impossible. For example, in Mudford and colleagues' (2011) calibration study described earlier, true values were obtained by a procedure that (a) was different from the procedure used by the observers to record their data (the researchers paused and replayed the video recordings while measuring the behaviors), and (b) incorporated extraordinary steps (the researcher resolved discrepant measures by viewing the video together until reaching agreement within 1-second tolerance).

It can be easy to mistake as true values, measures that only *appear* to be so. For example, suppose that four well-trained and experienced observers view a video recording of teacher and student interactions. Their task is to identify the true value of all instances of teacher praise contingent on academic accomplishments. Each observer views the recording independently and counts all occurrences of contingent teacher praise. After recording their respective observations, the four observers share their measurements, discuss disagreements, and suggest reasons for the disagreements. The observers independently record contingent praise a second time. Once again they share and discuss their results. After repeating the recording and sharing process several times, all observers agree that they have recorded every instance of teacher praise. However, the observers did not produce a *true value* of teacher praise for two reasons: (1) The observers could not calibrate their measurement of teacher praise to an independent standard of teacher praise, and (2) the process used to identify all instances of teacher praise may be biased (e.g., one of the observers may have convinced the others that her measures represented the true value). When true values cannot be established, researchers must rely on reliability assessments and measures of interobserver agreement to evaluate the quality of their data.

Accuracy Assessment Procedures

Determining the accuracy of measurement is a straightforward process of calculating the correspondence of each measure, or datum, assessed to its true value. For example, a researcher or practitioner assessing the accuracy of the score for a student's performance on a 30-word spelling test reported by a grader would compare the grader's scoring of each word on the test with the true value for that word found in a dictionary. Each word on the test that matched the correct letter sequence (i.e., orthography) provided by the dictionary and was marked correct by the grader would be an accurate measure by the grader, as would each word marked incorrect by the grader that did not match the dictionary's spelling. If the original grader's scoring of 29 of the test's 30 words corresponded to the true values for those words, the grader's measure would be 96.7% accurate.

Although an individual researcher or practitioner can assess the accuracy of the data she has collected, multiple independent observers are often used. Brown, Dunne, and Cooper (1996) described the procedures they used to assess the accuracy of measurement in a study of oral reading comprehension as follows:

> An independent observer reviewed one student's audio-tape of the delayed one-minute oral retell each day to assess our accuracy of measurement, providing an assessment of the extent that our counts of delayed retells approximated the true value of the audio-taped correct and incorrect retells. The independent observer randomly selected each day's audiotape by drawing a student's name from a hat, then listened to the tape and scored correct and incorrect retells using the same definitions as the teacher. Observer scores were compared to teacher scores. If there was a discrepancy between these scores, the observer and the teacher reviewed the tape (i.e., the true value) together to identify the source of the discrepancy and corrected the counting error on the data sheet and the Standard Celeration Chart. The observer also used a stopwatch to time the duration of the audio-tape

> to ensure accuracy of the timings. We planned to have the teacher re-time the presentation or retell and recalculate the frequency per minute for each timing discrepancy of more than 5 seconds. All timings, however, met the 5-second accuracy definition. (p. 392)

Reporting Accuracy Assessments

In addition to describing procedures used to assess the accuracy of the data, researchers should report the number and percentage of measures that were checked for accuracy, the degree of accuracy found, the extent of measurement error detected, and whether those measurement errors were corrected in the data. Brown and colleagues (1996) used the following narrative to report the results of their accuracy assessment:

> The independent observer and the teacher achieved 100% agreement on 23 of the 37 sessions checked. The teacher and the observer reviewed the tape together to identify the source of measurement errors for the 14 sessions containing measurement discrepancies and corrected the measurement errors. Accurate data from the 37 sessions rechecked were then displayed on the Standard Celeration Charts. The magnitude of the measurement errors was very small, often a difference of 1 to 3 discrepancies. (p. 392)

A full description and reporting of the results of accuracy assessment helps readers of the study evaluate the accuracy of all the data included in the report. For example, suppose a researcher reported that she conducted accuracy checks on a randomly selected 20% of the data, found those measures to be 97% accurate with the 3% error being nonbiased, and corrected the assessed data as needed. A reader of the study would know that 20% of the data are 100% accurate and be fairly confident that the remaining 80% of the data (i.e., all of the measures that were not checked for accuracy) is 97% accurate.

Assessing the Reliability of Measurement

Measurement is reliable when it yields the same values across repeated measures of the same event. Reliability is established when the same observer measures the same data set repeatedly from archived response products such as audiovisual products and other forms of permanent products. The more frequently a consistent pattern of observation is produced, the more reliable the measurement. Conversely, if similar observed values are not achieved with repeated observations, the data are considered unreliable. This leads to a concern about accuracy, which is the primary indicator of quality measurement.

But, as we have pointed out repeatedly, reliable data are not necessarily accurate data. As the three bicyclists discovered, totally reliable (i.e., consistent) measurement may be totally wrong. Relying on the reliability of measurement as the basis for determining the accuracy of measurement would be, as the philosopher Wittgenstein (1953) noted, "As if someone were to buy several copies of the morning paper to assure himself that what it said was true" (p. 94).

In many research studies and most practical applications, however, checking the accuracy of every measure is not possible or feasible. In other cases, true values for measures of the target behavior may be difficult to establish. When confirming the accuracy of each datum is not possible or practical, or when true values are not available, knowing that a measurement system has been applied with a high degree of consistency contributes to confidence in the overall trustworthiness of the data. Although high reliability cannot confirm high accuracy, discovering a low level of reliability signals that the data are then suspect enough to be disregarded until problems in the measurement system can be determined and repaired.

Assessing the reliability of behavioral measurement requires either a natural or contrived permanent product so the observer can re-measure the same events. For example, reliability of measurement of variables such as the number of adjectives or action verbs in students' essays could be accomplished by having an observer rescore essays. The reliability of the measurement reporting the number and type of response prompts and feedback statements parents presented to their children at the family dinner table could be evaluated by having an observer replay and rescore videos of the family's mealtime and compare the data obtained from the two measurements.

Observers should not re-measure the same permanent product soon after measuring it the first time. Doing so might result in the measures from the second scoring being influenced by what the observer remembered from the initial scoring. To avoid such unwanted influence, a researcher can insert several previously scored essays or videos randomly into the sequence of "new data" being recorded by observers.

USING INTEROBSERVER AGREEMENT TO ASSESS BEHAVIORAL MEASUREMENT

Interobserver agreement is the most commonly used indicator of measurement quality in applied behavior analysis (Kostewicz et al., 2016). **Interobserver agreement (IOA)** refers to the degree to which two or more independent observers report the same observed values after measuring the same events.

Benefits and Uses of IOA

Obtaining and reporting interobserver agreement serves four distinct purposes. First, IOA can function as a basis for determining the competence of new observers. As noted earlier, a high degree of agreement between a newly trained observer and an experienced observer provides an objective index of the extent to which the new observer is measuring the behavior in the same way as experienced observers.

Second, systematic assessment of IOA over the course of a study can detect observer drift. When observers who obtained the same, or nearly the same, observed values when measuring the same behavioral events at the beginning of a study (i.e., IOA was high) obtain different measures of the same events later in the study (i.e., IOA is now low), one of the observers may be using a definition of the target behavior that has drifted. Deteriorating IOA assessments cannot indicate with assurance which of the observer's data are being influenced by drift (or any other reason for disagreement), but the information reveals the need for further evaluation of the data and/or for retraining and calibration of the observers.

Third, knowing that two or more observers consistently obtained similar data increases confidence that the definition of the target behavior was clear and unambiguous and the measurement code and system not too difficult. Fourth, for studies that employ multiple observers as data collectors, consistently high levels of IOA increase confidence that variability in the data is not a function of which observer(s) happened to be on duty for any given session, and therefore that changes in the data more likely reflect actual changes in the behavior.

The first two reasons for assessing IOA are proactive: They help researchers determine and describe the degree to which observers have met training criteria and detect possible drift in observers' use of the measurement system. The second two purposes or benefits of IOA are as summative descriptors of the consistency of measurement across observers. By reporting the results of IOA assessments, researchers enable consumers to judge the relative **believability** of the data as trustworthy and deserving of interpretation.

Requisites for Obtaining Valid IOA Measures

A valid assessment of IOA depends on three equally important criteria. Although these criteria are perhaps obvious, it is nonetheless important to make them explicit. Two observers (usually two, but there may be more) must (a) use the same observation code and measurement system, (b) observe and measure the same participant(s) and events, and (c) observe and record the behavior independent of any influence from one other.

Observers Must Use the Same Measurement System

Interobserver agreement assessments conducted for any of the four previously stated reasons require observers to use the same definitions of the target behavior, observation procedures and codes, and measurement devices. Beyond using the same measurement system, all observers participating in IOA measures used to assess the believability of data (as opposed to evaluating the observer trainees' performance) should have received identical training with the measurement system and achieved the same level of competence in using it.

Observers Must Measure the Same Events

The observers must observe the same subject(s) at precisely the same observation intervals and periods. IOA for data obtained by real-time measurement requires that both observers be in the setting simultaneously. Real-time observers must be positioned such that each has a similar view of the subject(s) and environment. Two observers sitting on opposite sides of a classroom, for example, might obtain different measures because the different vantage points enable only one observer to see or hear some occurrences of the target behavior.

Observers must begin and end the observation period at precisely the same time. Even a difference of a few seconds between observers may produce significant measurement disagreements. To remedy this situation, the timing devices could be started simultaneously and outside the observation setting, but before data collection begins, with the understanding that the data collection would actually start at a prearranged time (e.g., exactly at the beginning of the fifth minute). Alternatively, but less desirably, one observer could signal the other at the exact moment the observation is to begin.

A common and effective procedure is for both observers to listen by earphones to prerecorded audio cues signaling the beginning and end of each observation interval (see Chapter 4). An inexpensive splitter device that enables two earphones to be plugged into the same playback device allows observers to receive simultaneous cues unobtrusively and without depending on one another.

When assessing IOA for data obtained from permanent products, the two observers do not need to measure the behavior simultaneously. For example, the observers could each watch and record data from the same video or audio recording at different times. Procedures must be in place, however, to ensure that both observers watched or listened to the same recording and that they started and stopped their independent observations at precisely the same point(s) on the tapes. Ensuring that two observers measure the same events when the target behavior produces natural permanent products, such as completed academic assignments or widgets manufactured, would include procedures such as clearly marking the session number, date, condition, and subject's name on the product and guarding the response products to ensure they are not disturbed until the second observer has obtained his measure.

Observers Must Be Independent

The third essential ingredient for valid IOA assessment is ensuring that neither observer is influenced by the other's measurements. Procedures must be in place to guarantee each observer's independence. For example, observers conducting real-time measurement of behavior must not be seated or positioned so closely to one another that either observer can detect or be influenced by the other observer's recordings (Johnston & Pennypacker, 2009).

Giving the second observer academic worksheets or written assignments that have already been marked by another observer would violate the observers' independence. To maintain independence, the second observer must score photocopies of unadulterated and unmarked worksheets or assignments as completed by the subjects.

Methods for Calculating IOA

Numerous methods for calculating IOA are available, each of which provides a somewhat different view of the extent and nature of agreement and disagreement between observers (e.g., Hartmann, 1977; Hawkins & Dotson, 1975; Page & Iwata, 1986; Mudford, Taylor, & Martin, 2009; Repp, Dietz, Boles, Dietz, & Repp, 1976). The following explanation of different IOA formats is organized by the three major methods for measuring behavioral data described in Chapter 4: event recording, timing, and interval recording or time sampling. Although other statistics are sometimes used, the percentage of agreement between observers is by far the most common convention for reporting IOA in applied behavior analysis.[6] Therefore, we have provided the formula for calculating a percentage of agreement for each type of IOA.

The formula for calculating **mean duration-per-occurrence IOA** is similar to the one used to determine mean count-per-interval IOA:

$$\frac{\text{Dur IOA R1} + \text{Dur IOA R2} + \text{Dur IOA R}n}{n \text{ responses with Dur IOA}} \times 100 = \text{mean duration-per-interval IOA }\%$$

Using this formula to calculate the mean duration-per-occurrence IOA for the two observers' timing data of the five responses just presented would entail the following steps:

1. Calculate duration-per-occurrence IOA for each response: R1, 29 ÷ 0.35 = 0.83; R2, 15 ÷ 21 = 0.71; R3, 7 ÷ 9 = 0.78; R4, 14 ÷ 14 = 1.0; and R5, 14 ÷ 17 = 0.82
2. Add the individual IOA percentages for each occurrence: 0.83 + 0.71 + 0.78 + 1.00 + 0.82 = 4.14
3. Divide the sum of the individual IOAs per occurrence by the total number of responses for which two observers measured duration: 4.14 ÷ 5 = 0.828
4. Multiply by 100 and round to the nearest whole number: 0.828 × 100 = 83%

This basic formula is also used to compute the *mean latency-per-response IOA* or *mean IRT-per-response IOA* for latency and IRT data. An observer's timings of latencies or IRTs in a session should never be added and the total time compared to a similar total time obtained by another observer as the basis for calculating IOA for latency and IRT measures.

In addition to reporting mean agreement per occurrence, IOA assessment for timing data can be enhanced with information about the range of differences between observers' timings and the percentage of responses for which the two observers each obtained measures within a certain range of error. For example, a researcher might report: "Mean duration-per-occurrence IOA for Participant 1's compliance was 87% (range across responses, 63% to 100%), and 96% of all timings obtained by the second observer were within +/−2 seconds of the primary observer's measures."

IOA for Data Obtained by Interval Recording/Time Sampling

Three techniques commonly used by applied behavior analysts to calculate IOA for interval data are interval-by-interval IOA, scored-interval IOA, and unscored-interval IOA.

Interval-by-Interval IOA. When using an interval-by-interval IOA (sometimes referred to as the *point-by-point, total interval*, or *block-by-block method*), the primary observer's record for each interval is matched to the secondary observer's record for the same interval. The formula for calculating **interval-by-interval IOA** is as follows:

$$\frac{\text{Number of intervals agreed}}{\text{Number of intervals agreed} + \text{number of intervals disagreed}} \times 100 = \text{interval-by-interval IOA }\%$$

The hypothetical data in Figure 5.3 show the interval-by-interval method for calculating IOA based on the record of two observers who recorded the occurrence (X) and nonoccurrence (0) of behavior in each of 10 observation intervals. The observers' data sheets show that they agreed on the occurrence or the nonoccurrence of the behavior for seven intervals (Intervals 2, 3, 4, 5, 7, 9, and 10). Interval-by-interval IOA for this data set is 70% (i.e., 7 ÷ [7 + 3] × 100 = 70%).

Interval-by-interval IOA is likely to overestimate the actual agreement between observers measuring behaviors that occur at very low or very high rates. This is because interval-by-interval IOA is subject to random or accidental agreement between observers. For example, with a behavior whose actual frequency of occurrence is only about 1 or 2 intervals per 10 observation intervals, even a poorly trained and unreliable observer who misses some of the few occurrences of the behavior and mistakenly records the behavior as occurring in some intervals when the behavior did not occur is likely to mark most intervals as nonoccurrences. As a result of this chance agreement, interval-by-interval IOA is likely to be quite high. Two IOA methods that minimize the effects of chance agreements for interval data on behaviors that occur at very low or very high rates are scored-interval IOA and unscored-interval IOA (Hawkins & Dotson, 1975).

Figure 5.3 When calculating interval-by-interval IOA, the number of intervals in which both observers agreed on the occurrence or the nonoccurrence of the behavior (shaded intervals: 2, 3, 4, 5, 7, 9, 10) is divided by the total number of observation intervals (10). Interval-by-interval IOA for the data shown here is 70% (7/10).

Interval-by-Interval IOA

Interval no. →	1	2	3	4	5	6	7	8	9	10
Observer 1	X	X	X	0	X	X	0	X	X	0
Observer 2	0	X	X	0	X	0	0	0	X	0

X = behavior was recorded as occurring during interval
0 = behavior was recorded as not occurring during interval

Scored-Interval IOA. Only those intervals in which either or both observers recorded the *occurrence of the target behavior* are used in calculating **scored-interval IOA**. An agreement is counted when both observers recorded that the behavior occurred in the same interval, and each interval in which one observer recorded the occurrence of the behavior and the other recorded its nonoccurrence is counted as a disagreement. For example, for the data shown in Figure 5.4, only Intervals 1, 3, and 9 would be used in calculating scored-interval IOA. Intervals 2, 4, 5, 6, 7, 8, and 10 would be ignored because both observers recorded that the behavior did not occur in those intervals. Because the two observers agreed that the behavior occurred in only one (Interval 3) of the three scored intervals, the scored-interval IOA measure is 33% (1 interval of agreement divided by the sum of 1 interval of agreement plus 2 intervals of disagreement × 100 = 33%).

For behaviors that occur at low rates, scored-interval IOA is a more conservative measure of agreement than interval-by-interval IOA. This is because scored-interval IOA ignores the intervals in which agreement by chance is highly likely. For example, using the interval-by-interval method for calculating IOA for the data in Figure 5.4 would yield an agreement of 80%. To avoid overinflated and possibly misleading IOA measures, we recommend using scored-interval interobserver agreement for behaviors that occur at frequencies of approximately 30% or fewer intervals.

Unscored-Interval IOA. Only intervals in which either or both observers recorded the *nonoccurrence of the target behavior* are considered when calculating **unscored-interval IOA**. An agreement is counted when both observers recorded the nonoccurrence of the behavior in the same interval, and each interval in which one observer recorded the nonoccurrence of the behavior and the other recorded its occurrence is counted as a disagreement. For example, only Intervals 1, 4, 7, and 10 would be used in calculating the unscored-interval IOA for the data in Figure 5.5 because at least one observer recorded the nonoccurrence of the behavior in each of those intervals. The two observers agreed that the behavior did not occur in Intervals 4 and 7. Therefore, the unscored-interval IOA in this example is 50% (2 intervals of agreement divided by the sum of 2 intervals of agreement plus 2 intervals of disagreement × 100 = 50%).

For behaviors that occur at relatively high rates, unscored-interval IOA provides a more stringent assessment of interobserver agreement than does interval-by-interval IOA. To avoid overinflated and possibly misleading IOA measures, we recommend using unscored-interval interobserver agreement for behaviors that occur at frequencies of approximately 70% or more of intervals.

The equations for calculating each of the aforementioned types of IOA are summarized in Figure 5.6.

Considerations in Selecting, Obtaining, and Reporting Interobserver Agreement

The guidelines and recommendations that follow are organized under a series of questions concerning the use of interobserver agreement to evaluate the quality of behavioral measurement.

Figure 5.4 Scored-interval IOA is calculated using only those intervals in which either observer recorded the occurrence of the behavior (shaded intervals: 1, 3, 9). Scored-interval IOA for the data shown here is 33% (1/3: interval 3).

Scored-Interval IOA

Interval no. →	1	2	3	4	5	6	7	8	9	10
Observer 1	X	0	X	0	0	0	0	0	0	0
Observer 2	0	0	X	0	0	0	0	0	X	0

X = behavior was recorded as occurring during interval
0 = behavior was recorded as not occurring during interval

Figure 5.5 Unscored-interval IOA is calculated using only those intervals in which either observer recorded the nonoccurrence of the behavior (shaded intervals: 1, 4, 7, 10). Unscored-interval IOA for the data shown here is 50% (2/4: Interval 4 and 7).

Unscored-Interval IOA

Interval no. →	1	2	3	4	5	6	7	8	9	10
Observer 1	X	X	X	0	X	X	0	X	X	0
Observer 2	0	X	X	0	X	X	0	X	X	X

X = behavior was recorded as occurring during interval
0 = behavior was recorded as not occurring during interval

Figure 5.6 Equations for calculating various types of interobserver agreement (IOA).

Type of IOA	Equation
Total count	$\frac{\text{smaller count}}{\text{larger count}} \times 100$
Mean count-per-interval	$\frac{\text{Int 1 IOA} + \text{Int 2 IOA} + \text{Int } n \text{ IOA}}{n \text{ intervals}} \times 100$
Exact count-per-interval	$\frac{\text{number of intervals of 100\% IOA}}{n \text{ intervals}} \times 100$
Trial-by-trial	$\frac{\text{number of trials (items) agreed}}{\text{total number of trials (items)}} \times 100$
Total duration	$\frac{\text{shorter duration}}{\text{longer duration}} \times 100$
Mean duration-per-occurrence	$\frac{\text{Dur IOA R1} + \text{Dur IOA R2} + \text{Dur IOA R}n}{n \text{ responses with Dur IOA}} \times 100$
Interval-by-zinterval	$\frac{\text{number of intervals agreed}}{\text{no. intervals agreed} + \text{no. intervals disagreed}} \times 100$
Scored-interval	$\frac{\text{no. of intervals in which either observer recorded the occurrence of the behavior}}{\text{total number of all scored intervals}} \times 100$
Unscored-interval	$\frac{\text{no. of intervals in which either observer recorded the nonoccurrence of the behavior}}{\text{total number of all unscored intervals}} \times 100$

How Often and When Should IOA Be Obtained?

Interobserver agreement should be assessed during each condition and phase of a study and be distributed across days of the week, times of day, settings, and observers. Scheduling IOA assessments in this manner ensures that the results will provide a representative (i.e., valid) picture of all data obtained in a study. It is generally recommended that IOA be obtained for a minimum of 20% of a study's sessions, and preferably between 25% and 33% of sessions (Gast & Ledford, 2014; Kennedy, 2005; Poling, Methot, & LeSage, 1995).

In general, studies using data obtained with real-time measurement will have IOA assessed for a higher percentage of sessions than studies with data obtained from permanent products.

The frequency with which data should be assessed using interobserver agreement will vary depending on the complexity of the measurement code, the number and experience of observers, the number of conditions and phases, and the results of the IOA assessments themselves. More frequent IOA assessments are expected in studies that involve complex or new measurement systems, inexperienced observers, and numerous conditions and phases. If appropriately conservative methods for obtaining and calculating IOA reveal high levels of agreement early in a study, the number and proportion of sessions in which IOA is assessed may decrease as the study progresses. For instance, IOA assessment might be conducted in each session at the beginning of an analysis, and then reduced to a schedule of once per four or five sessions.

For What Variables Should IOA Be Obtained and Reported?

In general, researchers should obtain and report IOA at the same levels at which they report and discuss the results of their study. For example, a researcher analyzing the relative effects of two treatment conditions on two behaviors of four participants in two settings should report IOA outcomes on both behaviors for each participant separated by treatment condition and setting. This would enable consumers of the research to judge the relative believability of the data within each component of the experiment.

Which Method of Calculating IOA Should Be Used?

No method for calculating IOA is universally superior. Because each method shows systematic bias in various situations (Mudford, Martin, Hui, & Taylor, 2009), behavior analysts should select IOA method(s) best suited to the measurement system and resulting data set. As a general rule, more stringent

and conservative methods of calculating IOA should be used over methods that are likely to overestimate actual agreement as a result of chance. With event recording data, we recommend reporting overall IOA on a trial-by-trial or item-by-item basis, perhaps supplemented with separate IOA calculations for correct responses and incorrect responses. For data obtained by interval or time sampling measurement, we recommend supplementing interval-by-interval IOA with scored-interval IOA or unscored-interval IOA, depending on the relative frequency of the behavior. In situations in which the primary observer scores the target behavior as occurring in approximately 30% or fewer intervals, scored-interval IOA provides a conservative supplement to interval-by-interval IOA. Conversely, when the primary observer scores the target behavior as occurring in approximately 70% or more of the intervals, unscored-interval IOA should supplement interval-by-interval IOA. If the rate at which the target behavior occurs changes from very low to very high, or from very high to very low, across conditions or phases of a study, reporting both unscored-interval and scored-interval IOA may be warranted.

If in doubt about which form of IOA to report, calculating and presenting several variations will help readers make their own judgments regarding the believability of the data. However, if the acceptance of the data for interpretation or decision making rests on which formula for calculating IOA is chosen, serious concerns about the data's trustworthiness exist and must be addressed.

What Are Acceptable Levels of IOA?

Carefully collected and conservatively computed IOA assessments increasingly enhance the believability of a data set as agreement approaches 100%. The usual convention in applied behavior analysis is to expect independent observers to achieve a mean of no less than 80% agreement when using observational recording. However, as Kennedy (2005) pointed out, "There is no scientific justification for why 80% is necessary, only a long history of researchers using this percentage as a benchmark of acceptability and being successful in their research activities" (p. 120).

According to Kleinmann et al. (2009), "convention stipulates that mean [interobserver] agreement should equal or exceed 85%" (p. 474). However, various factors at work in a given situation may make an 85% criterion too low or too high (Kazdin, 2011). For example, 95% IOA on the number of words contained in student compositions should raise serious questions about the trustworthiness of the data. IOA at or very near 100% is needed to enhance the believability of count data obtained from permanent products. However, some analysts might accept data with a mean IOA below 80% for the simultaneous measurement of multiple behaviors by several subjects in a complex environment, especially if it is based on a sufficient number of individual IOA assessments with a small range (e.g., 75% to 83%).

The degree of behavior change revealed by the data should also be considered when determining an acceptable level of interobserver agreement. When behavior change from one condition to another is small, the variability in the data might represent inconsistent observation more than actual change in the behavior. Therefore, the smaller the change in behavior across conditions, the higher the criterion should be for an acceptable IOA percentage.

How Should IOA Be Reported?

IOA scores can be reported in narrative, table, and graphic form. Whichever format is chosen, it is important to note how, when, and how often interobserver agreement was assessed.

Narrative Description. The most common approach for reporting IOA is a simple narrative description of the mean and range of agreement percentages. For example, Craft, Alber, and Heward (1998) described the methods and results of IOA assessments in a study in which four dependent variables were measured as follows:

> *Student recruiting and teacher praise.* A second observer was present for 12 (30%) of the study's 40 sessions. The two observers independently and simultaneously observed the 4 students, recording the number of recruiting responses they emitted and teacher praise they received. Descriptive narrative notes recorded by the observers enabled each recruiting episode to be identified for agreement purposes. Interobserver agreement was calculated on an episode-by-episode basis by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100%. Agreement for frequency of student recruiting ranged across students from 88.2% to 100%; agreement for frequency of recruited teacher praise was 100% for all 4 students; agreement for frequency of nonrecruited teacher praise ranged from 93.3% to 100%.
>
> *Academic work completion and accuracy.* A second observer independently recorded each student's work completion and accuracy for 10 (25%) sessions. Interobserver agreement for both completion and accuracy on the spelling worksheets was 100% for all 4 students. (p. 403)

Table. An example of reporting interobserver agreement outcomes in table format is shown in Table 5.1. Krantz and McClannahan (2003) reported the range and mean IOA computed for three types of social interactions by three children across each experimental condition.

Graphic Display. Interobserver agreement can be represented visually by plotting the measures obtained by the secondary observer on a graph of the primary observer's data, as shown in Figure 5.7. Looking at both observers' data on the same graph reveals the extent of agreement between the observers and the existence of observer drift or bias. The absence of observer drift is suggested in the hypothetical study shown in Figure 5.6 because the secondary observer's measures changed in concert with the primary observer's measures. Although the two observers obtained the same measure on only 2 of the 10 sessions in which IOA was assessed

TABLE 5.1 Example of Reporting Interobserver Agreement Results for Each Dependent Variable by Participant and Experimental Condition

Range and Mean Percentage Interobserver Agreement on Scripted Interaction, Elaborations, and Unscripted Interaction by Child and Condition

	Condition									
	Baseline		**Teaching**		**New recipient**		**Script fading**		**New activities**	
Type of interaction	**Range**	**M**	**Range**	**M**	**Range**	**M**	**Range**	**M**	**Range**	**M**
Scripted										
David			88–100	94		100		100		
Jeremiah			89–100	98		100		—[a]		
Ben			80–100	98		90		—[a]		
Elaborations										
David			75–100	95	87–88	88	90–100	95		
Jeremiah			83–100	95	92–100	96		—[a]		
Ben			75–100	95		95		—[a]		
Unscripted										
David		100		100	87–88	88	97–100	98	98–100	99
Jeremiah		100		100	88–100	94	93–100	96		98
Ben		100		100		100	92–93	92	98–100	99

[a]No data are available for scripted responses and elaborations in the script-fading condition, because interobserver agreement was obtained after scripts were removed (i.e., because scripts were absent, there could be only unscripted responses).

From "Social Interaction Skills for Children with Autism: A Script-Fading Procedure for Beginning Readers," by P. J. Krantz and L. E. McClannahan, 1998, *Journal of Applied Behavior Analysis, 31*, p. 196. Copyright 1998 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

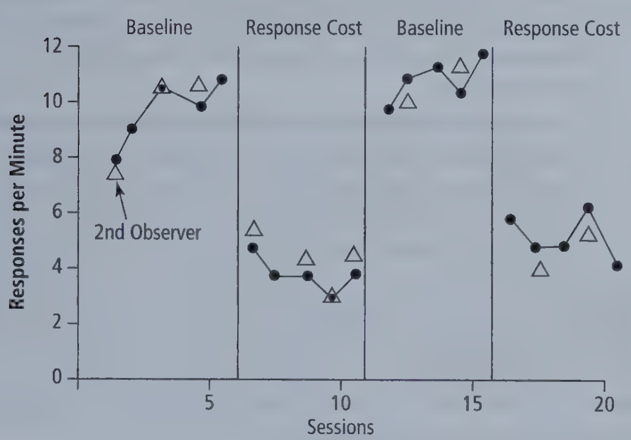


Figure 5.7 Plotting the second observer's data on a graph with the primary observer's data provides a visual representation of the extent and nature of interobserver agreement and may enable assessment of observer drift, bias, and conclusions regarding the presence of a functional relation.

(Sessions 3 and 8), the fact that neither observer consistently reported measures that were higher or lower than the other suggests the absence of observer bias. An absence of bias is usually indicated by a random pattern of overestimation and underestimation.

In addition to revealing observer drift and bias, a third way that graphically displaying IOA assessments can enhance the believability of measurement is illustrated by the data in Figure 5.6. When the data reported by the primary observer show clear change in the behavior between conditions or phases and all of the measures reported by the secondary observer within each phase fall within the range of observed values obtained by the primary observer, confidence increases that the data represent actual changes in the behavior measured rather than changes in the primary observer's behavior due to drift or extra-experimental contingencies. Likewise, graphic displays that reveal disagreement among observers can reduce viewers' confidence that a functional relationship is evident (Ledford, Artman, Wolery, & Wehby, 2012; Ledford & Wolery, 2013).

Although published research reports in applied behavior analysis seldom include graphic displays of IOA,[8] measures in creating and using such displays during a study constitute a simple and direct way for researchers to detect patterns in the consistency (or inconsistency) with which observers are measuring behavior that might not be as evident in comparing a series of percentages.

Which Approach Should Be Used for Assessing the Quality of Measurement: Accuracy, Reliability, or Interobserver Agreement?

Assessments of the accuracy of measurement, the reliability of measurement, and the extent to which different observers obtain the same measures each provide different indications

of data quality. Ultimately, the reason for conducting any type of assessment of measurement quality is to obtain quantitative evidence that can be used for the dual purposes of improving measurement during the course of an investigation and judging and convincing others of the trustworthiness of the data.

After ensuring the validity of what they are measuring and how they are measuring it, applied behavior analysts should choose to assess the accuracy of measurement whenever possible, rather than reliability or interobserver agreement. If it can be determined that all measurements in a data set meet an acceptable accuracy criterion, questions regarding the reliability of measurement and interobserver agreement are moot. For data confirmed to be accurate, conducting additional assessments of reliability or IOA is unnecessary.

When assessing the accuracy of measurement is not possible because true values are unavailable, an assessment of reliability provides the next best quality indicator. If natural or contrived permanent products can be archived, applied behavior analysts can assess the reliability of measurement, allowing consumers to know that observers have measured behavior consistently from session to session, condition to condition, and phase to phase.

When true values and permanent product archives are unavailable, interobserver agreement provides a level of believability for the data. Although IOA is not a direct indicator of the validity, accuracy, or reliability of measurement, it has proven to be a valuable and useful research tool in applied behavior analysis. Reporting interobserver agreement has been an expected and required component of published research in applied behavior analysis for several decades. In spite of its limitations, "the homely measures of observer agreement so widely used in the field are exactly relevant" (Baer, 1977, p. 119) to efforts to develop a robust technology of behavior change.

> Percentage of agreement, in the interval-recording paradigm, does have a direct and useful meaning: how often do two observers watching one subject, and equipped with the same definitions of behavior, see it occurring or not occurring at the same standard times? The two answers, "They agree about its occurrence X% of the relevant intervals, and about its nonoccurrence Y% of the relevant intervals," are superbly useful. (Baer, 1977, p. 118)

There are no reasons to prevent researchers from using multiple assessment procedures to evaluate the same data set. When time and resources permit, a combination of data assessments may be desirable. Applied behavior analysts can use any possible combination assessment (e.g., accuracy plus IOA, reliability plus IOA). In addition, some aspects of the data set could be assessed for accuracy or reliability, while other aspects are assessed with IOA. The previous example of accuracy assessment reported by Brown and colleagues (1996) included assessments for accuracy and IOA. Independent observers recorded correct and incorrect student-delayed retells. When IOA was less than 100%, data for that student and session were assessed for accuracy. IOA was used as an assessment to enhance believability, and also as a procedure for selecting data to be assessed for accuracy.

SUMMARY

Indicators of Trustworthy Measurement

1. To be most useful for science, measurement must be valid, accurate, and reliable.
2. Valid measurement in ABA encompasses three equally important elements: (a) measuring directly a socially significant target behavior, (b) measuring a dimension of the target behavior relevant to the question or concern about the behavior, and (c) ensuring that the data are representative of the behavior under conditions and during times most relevant to the reason(s) for measuring it.
3. Measurement is accurate when observed values, the data produced by measuring an event, match the true state, or true values, of the event.
4. Measurement is reliable when it yields the same values across repeated measurement of the same event.

Threats to Valid Measurement Validity

5. Indirect measurement—measuring a behavior different from the behavior of interest—threatens validity because it requires the researcher or practitioner to make inferences about the relationship between the measures obtained and the actual behavior of interest.
6. A researcher who uses indirect measurement must provide evidence that the behavior measured directly reflects, in some reliable and meaningful way, something about the behavior for which the researcher wishes to draw conclusions.
7. Measuring a dimension of the behavior that is ill suited for, or irrelevant to, the reason for measuring the behavior compromises validity.
8. Measurement artifacts are data that give an unwarranted or misleading picture of the behavior because of the way measurement was conducted. Discontinuous measurement, poorly scheduled observations, and insensitive or limiting measurement scales are common causes of measurement artifacts.

Threats to Accurate and Reliable Measurement

9. Most investigations in applied behavior analysis use human observers to measure behavior, and human error is the biggest threat to the accuracy and reliability of data.
10. Factors that contribute to measurement error include poorly designed measurement systems, inadequate observer training, and expectations about what the data should look like.

11. Observers should receive systematic training and practice with the measurement system and meet predetermined accuracy and reliability criteria before collecting data.
12. Observer drift—unintended changes in the way an observer uses a measurement system over the course of an investigation—can be minimized by booster training sessions and feedback on the accuracy and reliability of measurement.
13. An observer's expectations or knowledge about predicted or desired results can impair the accuracy and reliability of data.
14. Observers should not receive feedback about the extent to which their data confirm or run counter to hypothesized results or treatment goals.
15. Measurement bias caused by observer expectations can be avoided by using naive observers.
16. Observer reactivity is measurement error caused by an observer's awareness that others are evaluating the data he reports.

Assessing the Accuracy and Reliability of Behavioral Measurement

17. Researchers and practitioners who assess the accuracy of their data can (a) determine early in an analysis whether the data are usable for making experimental or treatment decisions, (b) discover and correct measurement errors, (c) detect consistent patterns of measurement error that can lead to the overall improvement or calibration of the measurement system, and (d) communicate to others the relative trustworthiness of the data.
18. Assessing the accuracy of measurement is a straightforward process of calculating the correspondence of each measure, or datum, assessed to its true value.
19. True values for many behaviors of interest to applied behavior analysts are evident and universally accepted or can be established conditionally by local context. True values for some behaviors (e.g., cooperative play) are difficult because the process for determining a true value must be different from the measurement procedures used to obtain the data one wishes to compare to the true value.
20. Assessing the extent to which observers are reliably applying a valid and accurate measurement system provides a useful indicator of the overall trustworthiness of the data.
21. Assessing the reliability of measurement requires a natural or contrived permanent product so the observer can remeasure the same behavioral events.
22. Although high reliability does not confirm high accuracy, discovering a low level of reliability signals that the data are suspect enough to be disregarded until problems in the measurement system can be determined and repaired.

Using Interobserver Agreement to Assess Behavioral Measurement

23. The most commonly used indicator of measurement quality in ABA is interobserver agreement (IOA), the degree to which two or more independent observers report the same observed values after measuring the same events.
24. Researchers and practitioners use measures of IOA to (a) determine the competence of new observers, (b) detect observer drift, (c) judge whether the definition of the target behavior is clear and the system not too difficult to use, and (d) convince others of the relative believability of the data.
25. Measuring IOA requires that two or more observers (a) use the same observation code and measurement system, (b) observe and measure the same participant(s) and events, and (c) observe and record the behavior independent of influence by other observers.
26. There are numerous techniques for calculating IOA, each of which provides a somewhat different view of the extent and nature of agreement and disagreement between observers.
27. Percentage of agreement between observers is the most common convention for reporting IOA in ABA.
28. IOA for data obtained by event recording can be calculated by comparing (a) the total count recorded by each observer per measurement period, (b) the counts tallied by each observer during each of a series of smaller intervals of time within the measurement period, or (c) each observer's count of 1 or 0 on a trial-by-trial basis.
29. Total count IOA is the simplest and crudest indicator of IOA for event recording data, and exact count-per-interval IOA is the most stringent for most data sets obtained by event recording.
30. IOA for data obtained by timing duration, response latency, or interresponse time (IRT) is calculated in essentially the same ways as for event recording data.
31. Total duration IOA is computed by dividing the shorter of the two durations reported by the observers by the longer duration. Mean duration-per-occurrence IOA is a more conservative and usually more meaningful assessment of IOA for total duration data and should always be calculated for duration-per-occurrence data.
32. Three techniques commonly used to calculate IOA for interval data are interval-by-interval IOA, scored-interval IOA, and unscored-interval IOA.
33. Because it is subject to random or accidental agreement between observers, interval-by-interval IOA is likely to overestimate the degree of agreement between observers measuring behaviors that occur at very low or very high rates.

34. Scored-interval IOA is recommended for behaviors that occur at relatively low frequencies; unscored-interval IOA is recommended for behaviors that occur at relatively high frequencies.

35. IOA assessments should occur during each condition and phase of a study and be distributed across days of the week, times of day, settings, and observers.

36. Researchers should obtain and report IOA at the same levels at which they report and discuss the results of their study.

37. More stringent and conservative IOA methods should be used over methods that may overestimate agreement as a result of chance.

38. The convention for acceptable IOA has been a minimum of 80%, but there can be no set criterion. The nature of the behavior being measured and the degree of behavior change revealed by the data must be considered when determining an acceptable level of IOA.

39. IOA scores can be reported in narrative, table, and graphic form.

40. Researchers can use multiple indices to assess the quality of their data (e.g., accuracy plus IOA, reliability plus IOA).

KEY TERMS

accuracy
believability
calibration
continuous measurement
direct measurement
discontinuous measurement
exact count-per-interval IOA
indirect measurement
interobserver agreement (IOA)
interval-by-interval IOA
mean count-per-interval IOA
mean duration-per-occurrence IOA
measurement bias
naïve observer
observed value
observer drift
observer reactivity
reliability
scored-interval IOA
total count IOA
total duration IOA
trial-by-trial IOA
true value
unscored-interval IOA
validity

MULTIPLE-CHOICE QUESTIONS

1. A(n) _______ measure of the actual behavior of interest will always possess more validity than a(n) _______ measure.
   - **a.** Indirect, direct
   - **b.** Direct, indirect
   - **c.** Discontinuous, continuous
   - **d.** Reliable, accurate

   Hint: (See "Indicators of Trustworthy Measurement")

2. ________ measurement is consistent measurement.
   - **a.** Reliable
   - **b.** Valid
   - **c.** Accurate
   - **d.** Discontinuous

   Hint: (See "Indicators of Trustworthy Measurement")

3. When data give an unwarranted or misleading picture of the behavior because of the way measurement was conducted, the data are called a(n):
   - **a.** Bias
   - **b.** Artifact
   - **c.** Drift
   - **d.** Variable

   Hint: (See "Threats to Valid Measurement")

4. In general, artifacts can be limited by scheduling ________ observations.
   - **a.** Indirect
   - **b.** Infrequent
   - **c.** Frequent
   - **d.** Mini

   Hint: (See "Threats to Valid Measurement")

5. Cumbersome and difficult to use measurement ________ can lead to inaccurate and unreliable measurements.
   - **a.** Agreement
   - **b.** Instruments
   - **c.** Observers
   - **d.** Rulers

   Hint: (See "Threats to Accurate and Reliable Measurement")

6. Which of the following is not a suggested training method for potential observers?
   a. Select observers carefully
   b. Train observers to an objective standard of competency
   c. Use a systematic approach to training
   d. Explicit feedback about data collection as it relates to the hypothesis

   Hint: (See "Threats to Accurate and Reliable Measurement")

7. Reliable data are not necessarily _______ data.
   a. Complete
   b. Direct
   c. Accurate
   d. Real

   Hint: (See "Assessing the Accuracy and Reliability of Behavioral Measurement")

8. Which of the following is not a requisite for obtaining valid IOA measures?
   a. Use the same observation code and measurement system
   b. Observe and measure the same participant(s) and events
   c. Observe and record the behavior independent of any influence from one another
   d. Score all IOA while sitting side-by-side with another observer

   Hint: (See "Assessing the Accuracy and Reliability of Behavioral Measurement")

9. Which of the following is a limitation of total count IOA?
   a. Provides no assurance that the two observers recorded the same instances of behavior
   b. Useful for event recording
   c. Is expressed as percentage of agreement
   d. Is calculated by dividing the smaller of the counts by the larger count and multiplying by 100

   Hint: (See "Methods for Calculating IOA")

10. Which of the following is the most stringent method for calculating IOA for behaviors that occur at relatively high rates and obtained by interval recording?
    a. Total Count IOA
    b. Unscored-interval IOA
    c. Scored-interval IOA
    d. Mean Count-per-Interval IOA

    Hint: (See "Methods for Calculating IOA")

ESSAY-TYPE QUESTIONS

1. Discuss the difference between accuracy and reliability as they relate to behavioral data and suggest methods to evaluate both.

   Hint: (See "Threats to Valid Measurement")

2. Describe threats to measurement validity in applied behavior analysis. Suggest ways to control the threats you describe.

   Hint: (See "Threats to Valid Measurement")

3. Describe ways in which applied behavior analysis researchers can control human error before, during, and after data collection.

   Hint: (See "Threats to Accurate and Reliable Measurement")

4. Discuss a training strategy you might use for data collectors/observers.

   Hint: (See "Threats to Accurate and Reliable Measurement")

5. Discuss the role of collecting interobserver agreement measures in applied behavior analysis research. List various methods for assessing agreement for event recording and interval recordings of behavioral data.

   Hint: (See "Using Interobserver Agreement to Assess Behavioral Measurement")

NOTES

1. Strategies for increasing the accuracy of self-reports can be found in Critchfield, Tucker, and Vuchinich (1998) and Finney, Putnam, and Boyd (1998).
2. Single measures, such as pretests and posttests, can provide valuable information on a person's knowledge and skills before and after instruction or treatment. The use of *probes*, occasional but systematic measures, to assess maintenance and generalization of behavior change is discussed in Chapter 30.
3. We recommend using automatic data-recording devices whenever possible. For example, to measure the amount of exercise by boys on stationary bicycles, DeLuca and Holborn (1992) used magnetic counters that automatically recorded the number of wheel revolutions.
4. The official time in the United States can be accessed through the National Institute of Standards and Technology: https://www.time.gov/. The ytterbium clock, though experimental at this time, exceeds the accuracy of the cesium clock with respect to emerging standards of uncertainty, stability, and reproducibility!
5. The preferred spelling of a word may change (e.g., *judgement* becomes *judgment*), but in such cases a new true value is established.

6. IOA can be calculated by product-moment correlations, which range from +1.0 to −1.0. However, expressing IOA by correlation coefficients has two major weaknesses: (a) High coefficients can be achieved if one observer consistently records more occurrences of the behavior than the other, and (b) correlation coefficients provide no assurance that the observers agreed on the occurrence of any given instance of behavior (Poling, Methot, & LeSage, 1995). Hartmann (1977) described the use of *kappa (k)* as a measure of IOA. The *k* statistic was developed by Cohen (1960) as a procedure for determining the proportion of agreements between observers that would be expected as a result of chance. However, the *k* statistic is seldom reported in the behavior analysis literature.
7. Multiple terms are used in the applied behavior analysis literature for the same methods of calculating IOA, and the same terms are sometimes used with different meanings. We believe the IOA terms used here represent the discipline's most used conventions. To point out and preserve some meaningful distinctions among variations of IOA measures, we have introduced several terms.
8. A hand search by Artman, Wolery, and Yoder (2012) of all publications in the *Journal of Applied Behavior Analysis* from 1977 through May 2009 found just four studies that presented both observers' data on the same graph (Luce, Delquardi, & Hall, 1980; Tertinger, Greene, & Lutzker, 1984; Van Houten & Nau, 1980; Van Houten & Rolider, 1984). Artman et al. discovered no articles containing graphic displays of IOA published in *JABA* in the 25 years following the two instances in 1984.

PART 3

Evaluating and Analyzing Behavior Change

Part 2 described considerations and procedures for selecting and defining target behaviors; outlined detailed methods for measuring behavior; and examined techniques for improving, assessing, and reporting the veracity of behavioral measurement. The product of these measurements—*data*—is the medium with which behavior analysts work. But what does the behavior analyst do with data? The five chapters in Part 3 focus on the presentation and interpretation of behavioral data and the design, conduct, and evaluation of experiments that analyze the effects of interventions.

Chapter 6 presents the purpose and benefits of graphic displays of behavioral data, identifies the types of graphs used in applied behavior analysis, explains how to construct line graphs, and describes how behavior analysts interpret graphically displayed data.

Although measurement and graphic displays can reveal whether, when, and to what extent behavior has changed, they alone cannot reveal what brought about the behavior change. Chapters 7 through 10 are devoted to what constitutes *analysis* in applied behavior analysis. Chapter 7 outlines the concepts and assumptions underlining the analysis of behavior; describes the requisite components of an experiment in behavior analysis; and explains how researchers and practitioners apply steady state strategy and the three elements of basic logic—prediction, verification, and replication—to seek and verify functional relations between behavior and its controlling variables. Chapters 8 and 9 describe the logic and operation of reversal, multielement, multiple baseline, and changing criterion designs—the most commonly used experimental designs in applied behavior analysis. Chapter 10 covers a wide range of topics requisite to a more complete understanding of behavioral research. Beginning with the assumption that the research methods of any science should reflect the characteristics of its subject matter, we examine the importance of analyzing the behavior of the individual client or research participant, discuss the value of flexibility in experimental design, identify common confounds to the internal validity of experiments, present methods for assessing the social validity of behavior changes and the treatments that accomplish them, and explain how replication is used to determine the external validity of research. Chapter 10 concludes with a series of issues and questions to consider in evaluating the "goodness" of a published applied behavior analysis study.

PART 3

Evaluating and Analyzing Behavior Change

Part 2 described considerations and procedures for selecting and defining target behaviors; outlined detailed methods for measuring behavior; and examined techniques for improving, assessing, and reporting the veracity of behavioral measurement. The product of these measurements—*data*—is the medium with which behavior analysts work. But what does the behavior analyst do with data? The five chapters in Part 3 focus on the presentation and interpretation of behavioral data and the design, conduct, and evaluation of experiments that analyze the effects of interventions.

Chapter 6 presents the purpose and benefits of graphic displays of behavioral data, identifies the types of graphs used in applied behavior analysis, explains how to construct line graphs, and describes how behavior analysts interpret graphically displayed data.

Although measurement and graphic displays can reveal whether, when, and to what extent behavior has changed, they alone cannot reveal what brought about the behavior change. Chapters 7 through 10 are devoted to what constitutes *analysis* in applied behavior analysis. Chapter 7 outlines the concepts and assumptions underlining the analysis of behavior; describes the requisite components of an experiment in behavior analysis; and explains how researchers and practitioners apply steady state strategy and the three elements of basic logic—prediction, verification, and replication—to seek and verify functional relations between behavior and its controlling variables. Chapters 8 and 9 describe the logic and operation of reversal, multielement, multiple baseline, and changing criterion designs—the most commonly used experimental designs in applied behavior analysis. Chapter 10 covers a wide range of topics requisite to a more complete understanding of behavioral research. Beginning with the assumption that the research methods of any science should reflect the characteristics of its subject matter, we examine the importance of analyzing the behavior of the individual client or research participant, discuss the value of flexibility in experimental design, identify common confounds to the internal validity of experiments, present methods for assessing the social validity of behavior changes and the treatments that accomplish them, and explain how replication is used to determine the external validity of research. Chapter 10 concludes with a series of issues and questions to consider in evaluating the "goodness" of a published applied behavior analysis study.

CHAPTER 6

Constructing and Interpreting Graphic Displays of Behavioral Data

LEARNING OBJECTIVES

- State the purpose and list the benefits of graphic displays of behavioral data.
- List and describe the types of graphs used in applied behavior analysis.
- Given a set of behavioral data, select the appropriate data display to communicate quantitative relations.
- Given an unlabeled line graph, provide the correct labels.
- Describe and state the purpose of semilogarithmic charts.
- Describe and discuss the proper use of a cumulative record.
- Describe and discuss the proper use of a scatterplot.
- Given a set of behavioral data, construct and label a line graph.
- Given a set of behavioral data, select the most appropriate graphic display.
- Given a graphic display of behavioral data, identify the presence or absence of variability, level, and trends.

Applied behavior analysts quantify behavior change with direct and repeated measurement. The product of these measurements, called **data**, is the medium with which applied behavior analysts work. In everyday usage, the word *data* refers to a wide variety of often imprecise and subjective information offered as facts. In scientific usage *data* means "the results of measurement, usually in quantified form" (Johnston & Pennypacker, 1993a, p. 365).[1]

Because behavior change is dynamic and ongoing, behavior analysts—practitioners and researchers—strive for direct and continuous contact with the target behavior. The data obtained throughout a behavior change program or research study provide the empirical basis for every important decision: continue with the current intervention or experimental condition, implement a different intervention or condition, or reinstitute a previous condition. But making valid and reliable decisions from raw data (a series of numbers) is difficult, if not impossible, and inefficient. Inspecting a long sequence of numbers will reveal only large and abrupt changes in performance, or no change at all, and important features of behavior change can easily be overlooked.

Consider the three data sets that follow, each one a series of numbers representing consecutive measures of different target behaviors. The first data set shows the number of responses emitted under two conditions (A and B):

Condition A	Condition B
120, 125, 115, 130,	114, 110, 115, 121,
126, 130, 123, 120,	110, 116, 107, 120,
120, 127	115, 112

The second data set comprises consecutive measures of the percentage of correct responses:

80, 82, 78, 85, 80, 90, 85, 85, 90, 92

The third data set consists of responses per minute measured on successive school days:

35, 42, 33, 30, 25, 38, 41, 35, 35, 32, 40, 35, 49, 33, 30

What do these numbers tell you? What conclusions can you draw from each data set? How long did it take you to reach your conclusions? How sure of them are you? What if the data sets contained many more measures to interpret? How likely is it that others interested in the behavior change program or research study would reach the same conclusions? How could these data be directly and effectively communicated to others? This chapter will show you how graphs can help answer these questions.

Graphs—relatively simple formats for visually displaying relationships among and between a series of measurements and relevant variables—help people "make sense" of quantitative information. Graphs are the major device with which applied behavior analysts organize, store, interpret, and communicate the results of their work. Figure 6.1 includes a graph for each of the three data sets presented previously. The top graph reveals a level of responding during Condition B lower than measured during Condition A. The middle graph shows an upward trend over time in correct responding. Evident in the bottom graph is a variable pattern of responding, characterized by higher response rates during the first part of each week and a decreasing trend toward the end of each week. The graphs in Figure 6.1 illustrate three fundamental properties of behavior change over time—level, trend, and variability—each of

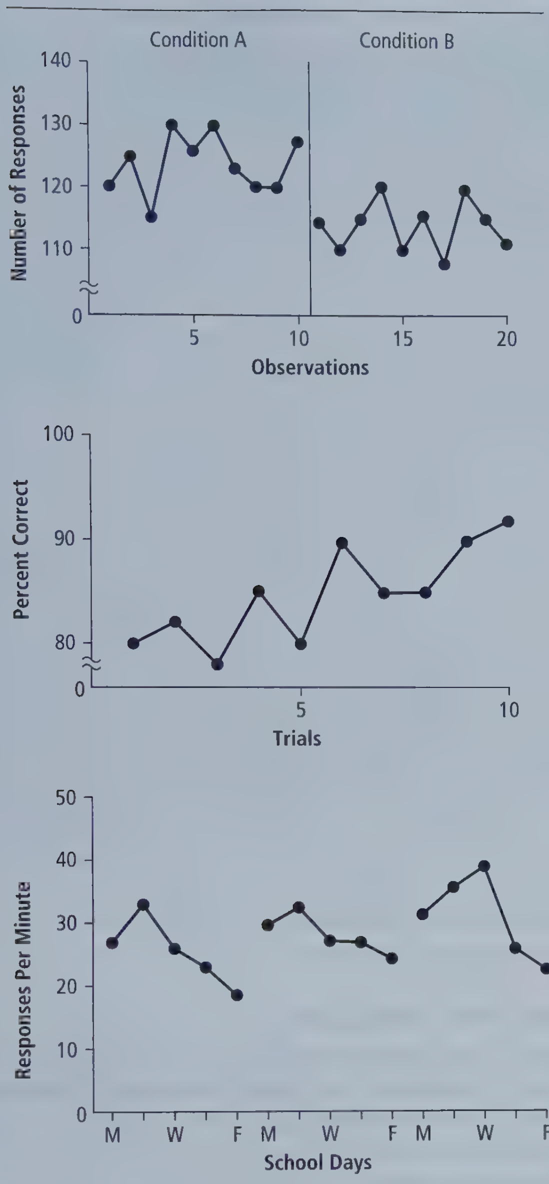


Figure 6.1 Graphic displays of three sets of hypothetical data illustrating changes in the level of responding across conditions (top), trend (middle), and cyclical variability (bottom).

which will be discussed in detail later in the chapter. Graphic displays of behavioral data are an effective means of illustrating and communicating these aspects of behavior change.

PURPOSE AND BENEFITS OF GRAPHIC DISPLAYS OF BEHAVIORAL DATA

> *In essence, the function of the graph is to communicate, in a readily assimilable and attractive manner, descriptions and summaries of data that enable rapid and accurate analysis of the facts.*
>
> —Parsonson and Baer (1978, p. 134)

The graphic display and visual analysis of behavioral data provide seven benefits to behavior analysts and their clients. First, plotting each measure of behavior on a graph soon after the observational period gives the practitioner or researcher immediate access to an ongoing visual record of the participant's behavior. Continual assessment of behavior change enables treatment and experimental decisions to be responsive to the participant's performance. Graphs provide the "close, continual contact with relevant outcome data" that can lead to "measurably superior instruction" (Bushell & Baer, 1994, p. 9).

Second, direct and continual contact with the data in a readily analyzable format enables the researchers to explore interesting variations in behavior as they occur. Some of the most important research findings about behavior have been made because scientists followed the leads suggested by their data instead of following predetermined experimental plans (Sidman, 1960, 1994; Skinner, 1956).

Third, graphs, like statistical analyses of behavior change, are judgmental aids: devices that help practitioners, researchers, and consumers interpret the results of a study or treatment (Michael, 1974). In contrast to the statistical tests of inference used in group comparison research, however, visual analysis of graphed data takes less time, is relatively easy to learn, imposes no predetermined or arbitrary level for determining the significance of behavior change, and does not require the data to meet preset mathematical properties or statistical assumptions to be analyzed.[2]

Fourth, visual analysis is a conservative method for determining the significance of behavior change. A behavior change deemed statistically significant by a test of mathematical probabilities may not look very impressive when the data are plotted on a graph that reveals great variability and repeated overlaps in the data within and across experimental or treatment conditions. This screening out of weak variables in favor of robust interventions has enabled applied behavior analysts to develop a useful technology of behavior change (Baer, 1977b).[3]

Fifth, graphs enable and encourage independent judgments and interpretations of the meaning and significance of behavior change. Instead of relying on conclusions based on statistical manipulations of the data or an author's interpretations, readers of published reports of applied behavior analysis can (and should) form independent conclusions based on their visual analysis of the data.[4]

Sixth, in addition to their primary purpose of displaying relationships between behavior change (or lack thereof) and variables manipulated by the researcher or practitioner, graphs can provide feedback to the people whose behavior the graphs depict. Numerous studies have reported improved performance of a wide variety of target behaviors by participants when receiving graphical feedback, alone or as part of a package intervention, (e.g., Gil & Carter, 2016; Grossi & Heward, 1998; Perrin, Fredrick, & Klick, 2016; Quinn, Miltenberger, Abreu, & Narozanick, 2017; Squires et al., 2008). Graphing one's own performance has also been demonstrated to be an effective intervention for a variety of academic and behavior change objectives (e.g., Kasper-Ferguson & Moxley, 2002; Stolz, Itoi, Konrad, & Alber-Morgan, 2008).

Seventh, graphs facilitate the communication, dissemination, and comprehension of behavior change among a wide variety of recipients (e.g., professionals, parents, government agents responsible for policymaking). A well-constructed

BOX 6.1

Page Space Devoted to Graphs by Behavior Analysis Journals

Open any issue of the *Journal of Applied Behavior Analysis (JABA)* to an article reporting an experimental study and chances are good you will see a graph. If not, turn a page or two in either direction, and a graph will almost certainly appear.

Does that matter? Yes, it does, for several reasons. First, *JABA's* frequent use of graphs emphasizes the central role of visual analysis of data in applied behavior analysis and underscores the obligation that practitioners, researchers, and journal article reviewers and editors know how to create and interpret graphic displays. Second, graphs take up a greater proportion of page space of journals in scientific disciplines (i.e., natural sciences) considered more mature, harder, and with greater consensus of knowledge than they do in less mature, softer scientific disciplines (i.e., social sciences) (Arsenault, Smith, & Beauchamp, 2006; Best, Smith, & Stubbs, 2001; Cleveland, 1984; Smith, Best, Stubbs, Johnston, & Bastiani-Archibald, 2000).

How do behavior analysis journals compare to other sciences in that regard? In an effort to shed light on that question, Kubina, Kostewicz, and Datchuk (2008) computed the fractional graph area (FGA)—the proportion of page space devoted to graphs—in research articles from one randomly selected issue from each of three volume years (1995, 2000, and 2005) from 11 behavioral journals. Diagrams and flow charts that did not graphically present quantitative data were not included.

Kubina and colleagues (2008) reported a mean FGA of .057 for all behavior analysis journals, which included journals such as *Behavior Therapy, Child and Family Therapy, Cognitive and Behavioral Practices*, and the *Journal of Behavior Therapy and Experimental Psychiatry* (see Figure A). The four behavior analysis journals with the highest FGA ranged from .08 for *The Analysis of Verbal Behavior* to .17 for the *Journal of the Experimental Analysis of Behavior (JEAB)*, equal to the FGAs of journals in the natural sciences, which range from .06 in geology to .18 in chemistry (as measured by Cleveland [1984]).

Notable among Kubina and colleagues' findings is that behavior analysis flagship journals' (*JABA* and *JEAB*) composite FGA of .147 ranked third among the natural sciences, trailing only chemistry and physics. While the authors cautioned that their results were obtained from just three issues from each journal, they noted that Best, Smith, and Stubbs's (2001) independent assessment of page space devoted to graphs yielded an identical measure of .17 for *JEAB*.

graph condenses aggregated individual numerical data so that relevant elements are visually salient and can be understood commonly.

Behavior analysts' heavy reliance on graphic data displays puts them in good company with scientists in other natural sciences (see Box 6.1).

GRAPHS USED BY APPLIED BEHAVIOR ANALYSTS

The graphic displays most often used by applied behavior analysts are line graphs, bar graphs, cumulative records, ratio charts, and scatterplots.

Line Graphs

The simple **line graph**, or frequency polygon, is far and away the most common graphic format in applied behavior analysis. Based on the Cartesian coordinate system created by René Descartes in the 17th century, the line graph is a two-dimensional area formed by the intersection of two perpendicular reference lines, called *axes*. Any point within the area enclosed by the two axes represents a specific relationship between values of the two dimensions (or variables) described by the intersecting lines. In applied behavior analysis, each point on a line graph shows the level of some quantifiable dimension of the target behavior (i.e., the **dependent variable**) in relation to a specified point in time and/or environmental condition (i.e., the **independent variable**) in effect when the measure was taken. Comparing points on a graph reveals the presence and extent of changes in level, trend, and/or variability within and across conditions.

Parts of a Basic Line Graph

Although graphs vary considerably in their final appearance, all properly constructed line graphs share certain elements. The basic parts of a simple line graph are shown in Figure 6.2 and described in the following sections.

1. Horizontal Axis. The *horizontal axis*, also called the *x*-axis, or *abscissa*, is a straight horizontal line that most often represents the passage of time and the presence, absence, and/or value of the independent variable. A defining characteristic of applied behavior analysis is the repeated measurement of behavior across time. Time is also the unavoidable dimension in which all manipulations of the independent variable occur. On most line graphs the passage of time is marked in equal intervals on the horizontal axis. The horizontal axis in Figure 6.2 represents consecutive observational sessions from 9:00 to 10:00 AM on weekdays.

The horizontal axis on some graphs represents different values of the independent variable instead of time. For example, Lalli, Mace, Livezey, and Kates (1998) scaled the horizontal axis of one graph in their study from less than 0.5 meter to 9.0 meters to show how the occurrence of self-injurious behavior by a girl with severe intellectual disabilities decreased as the distance between the therapist and the girl increased.

2. Vertical Axis. The *vertical axis*, also called the *y*-axis, or *ordinate*, is a vertical line drawn upward from the left-hand end of the horizontal axis. The vertical axis most often represents

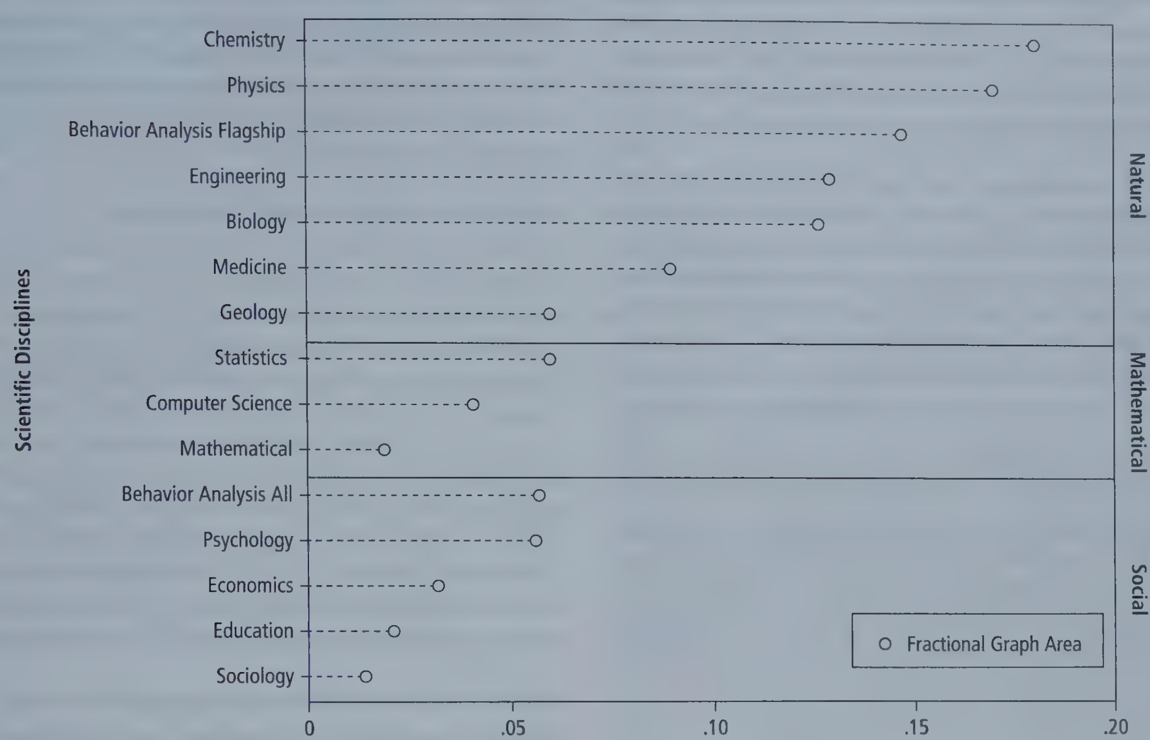


Figure A Proportion of page space devoted to graphic data displays by journals in various scientific disciplines. (Behavior analysis flagship journals are *JABA* and *JEAB*.)

From "An Initial Survey of Fractional Graph and Table Area in Behavioral Journals," by R. M. Kubina, D. E. Kostewicz, and S. M. D. Datchuk, 2008, *The Behavior Analyst, 31*, p. 65. Copyright 2008 by the Association for Behavior Analysis International. Used by permission.

a range of values of the dependent variable, which in applied behavior analysis is always some quantifiable dimension of behavior. The intersection of the horizontal and vertical axes is called the *origin* and most often, though not necessarily, represents the zero value of the dependent variable. Each successive point upward on the vertical axis represents a greater value of the dependent variable. The most common practice is to mark the vertical axis with an **equal-interval scale**, in which equal distances on the axis represent equal absolute amounts of behavior change. The vertical axis in Figure 6.2 represents the number of 50-meter laps walked in 1 hour.

3. Axis Labels. *Axis labels* are brief descriptions of the dimension represented by and scaled along each axis; most often a response measure on the vertical axis and experimental conditions across time on the horizontal axis.

4. Condition Change Lines. *Condition change lines* are vertical lines drawn upward from the horizontal axis to show points in time at which changes in the independent variable occurred. The condition change lines in Figure 6.2 coincide with the introduction or withdrawal of an intervention the researchers called Token Reinforcement.

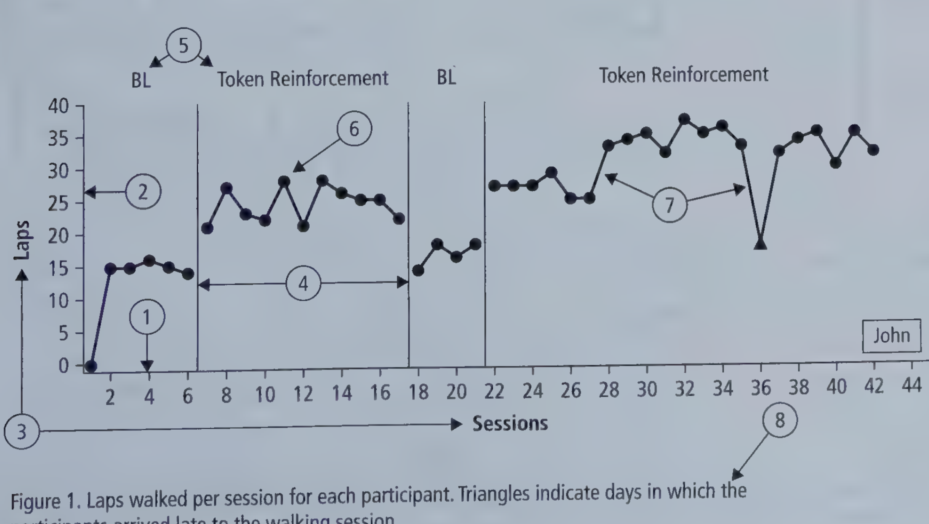


Figure 6.2 Major parts of a simple line graph: (1) horizontal axis, (2) vertical axis, (3) axis labels, (4) condition change lines, (5) condition labels, (6) data points, (7) data path, and (8) figure caption. (Original figure includes graphs for four additional participants.)

Based on "Using Token Reinforcement to Increase Walking for Adults with Intellectual Disabilities" by H. Krentz, R. Miltenberger, and D. Valbuena, 2016, *Journal of Applied Behavior Analysis, 49*, p. 749. Copyright 2016 by the Society for the Experimental Analysis of Behavior.

5. Condition Labels. *Condition labels* are single words or brief descriptive phrases printed along the top of the graph and parallel to the horizontal axis. These labels identify the experimental conditions (i.e., the presence, absence, or some value of the independent variable) in effect during each phase of the study.[5]

6. Data Points. Each *data point* on a graph represents two facts: (a) a quantifiable measure of the target behavior recorded during a given observation period and (b) the time and/or experimental conditions under which that particular measurement was conducted. Using two data points from Figure 6.2 as examples, we can see that John walked 14 laps during Session 6, the last session of the first baseline phase, and 21 laps in Session 7, the first session of the first phase of the token reinforcement condition.

7. Data Path. Connecting successive data points within a given condition with a straight line creates a **data path**. The data path represents the level and trend of behavior between successive data points and is a primary focus of attention in the interpretation and analysis of graphed data. Because behavior is rarely observed and recorded continuously in applied behavior analysis, the data path represents an estimate of the actual course taken by the behavior during the time between the two consecutive measures. The more measures and resultant data points per unit of time (given an accurate observation and recording system), the more confidence one can place in the story told by the data path.

8. Figure Caption. The *figure caption* is a concise statement that, in combination with the axis and condition labels, provides the reader with sufficient information to identify the independent and dependent variables. The figure caption should explain any symbols (see Figure 6.2) or observed but unplanned events that may have affected the dependent variable and point out and clarify any potentially confusing features of the graph (see Figure 6.6 and Figure 6.7).

Variations of the Simple Line Graph: Multiple Data Paths

The line graph is a remarkably versatile vehicle for displaying behavior change. Whereas Figure 6.2 is an example of the line graph in its simplest form (one data path showing a series of successive measures of behavior across time and experimental conditions), by the addition of multiple data paths, the line graph can display more complex behavior–environment relations. Graphs with multiple data paths are used frequently in applied behavior analysis to show (a) two or more dimensions of the same behavior, (b) two or more different behaviors, (c) the same behavior under different and alternating experimental conditions, (d) changes in target behavior relative to the changing values of an independent variable, and (e) the behavior of two or more participants.

Two or More Dimensions of the Same Behavior. Showing multiple dimensions of the dependent variable on the same graph enables visual analysis of the absolute and relative effects of the independent variable on those dimensions. Figure 6.3 shows the results of a study of the effects of training three members of a women's college basketball team proper foul shooting form

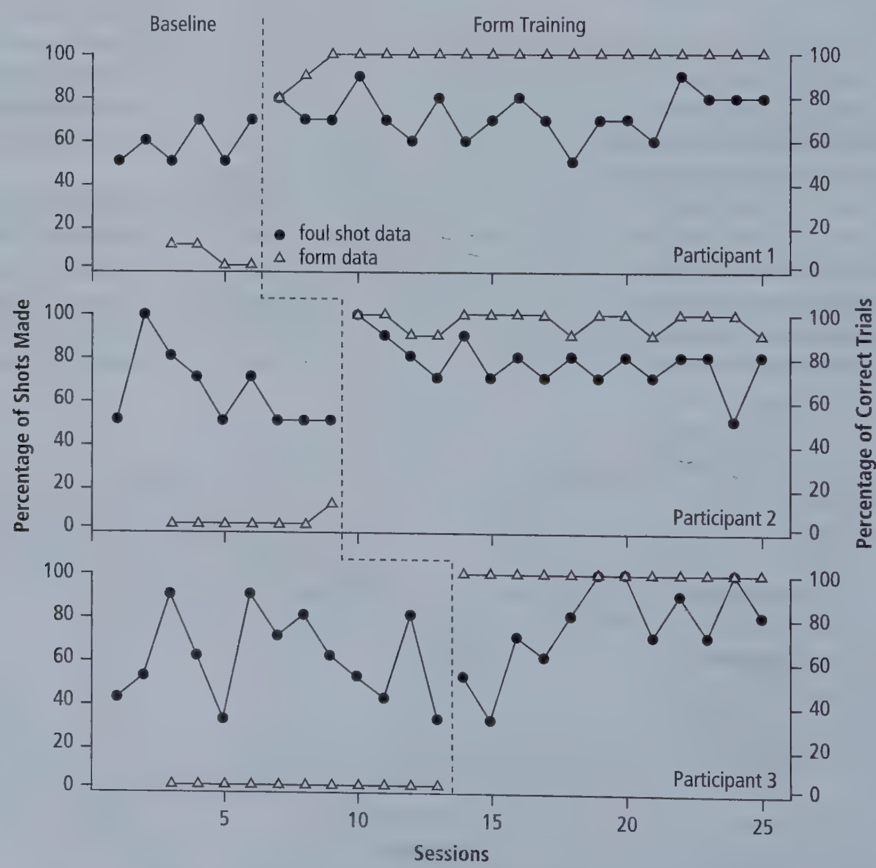


Figure 1. Percentage of shots made (filled circles) and percentage of shots taken with correct form (open triangles) across sessions for each participant.

Figure 6.3 Graph with multiple data paths showing the effects of the independent variable (Form Training) on two dimensions (accuracy and topography) of the dependent variable.

From "The Effects of Form Training on Foul-Shooting Performance in Members of a Women's College Basketball Team," by C. N. Kladopoulos and J. J. McComas, 2001, *Journal of Applied Behavior Analysis, 34*, p. 331. Copyright 2001 by the Society for the Experimental Analysis of Behavior, Inc. Used by permission.

(Kladopoulos & McComas, 2001). The data path created by connecting the open triangle data points shows changes in the percentage of foul shots executed with the proper form, whereas the data path connecting the solid data points reveals the percentage of foul shots made. Had the experimenters recorded and graphed only the players' foul shooting form, they would not have known whether any improvements in the target behavior on which training was focused (correct foul shooting form) coincided with improvements in the behavior by which the social significance of the study would ultimately be judged—foul shooting accuracy. By measuring and plotting both form and outcome on the same graph, the experimenters were able to analyze the effects of their treatment procedures on two critical dimensions of the dependent variable.

Two or More Different Behaviors. Multiple data paths are also used to facilitate simultaneous comparison of the effects of experimental manipulations on two or more different behaviors. Determining the variation of two behaviors as a function of changes in the independent variable is accomplished more easily if both behaviors can be displayed on the same graph. Figure 6.4 shows the percentage of intervals in which a boy with autism exhibited stereotypy (e.g., repetitive body movements, rocking) across three conditions, and the number of times that he raised his hand for attention (in the attention condition), signed for a break (in the demand condition), and signed for access to preferred tangible stimuli (in the no-attention condition) in a study investigating a strategy called *functional communication training* (Kennedy, Meyer, Knowles, & Shukla, 2000).[6] By recording and graphing both stereotypic responding and appropriate behavior, the investigators were able to determine whether increases in alternative communication responses (raising his hand and signing) were accompanied by reductions in stereotypy. Note that Figure 6.4 includes a second vertical axis showing the dimensional units and scaling for signing frequency. Because of the differences in scale, readers of dual-vertical axis graphs must view them with care, particularly when assessing the magnitude of behavior change.

Measures of the Same Behavior Under Different Conditions. Multiple data paths can also be used to represent measures of the same behavior taken under different experimental conditions that alternate throughout an experimental phase.

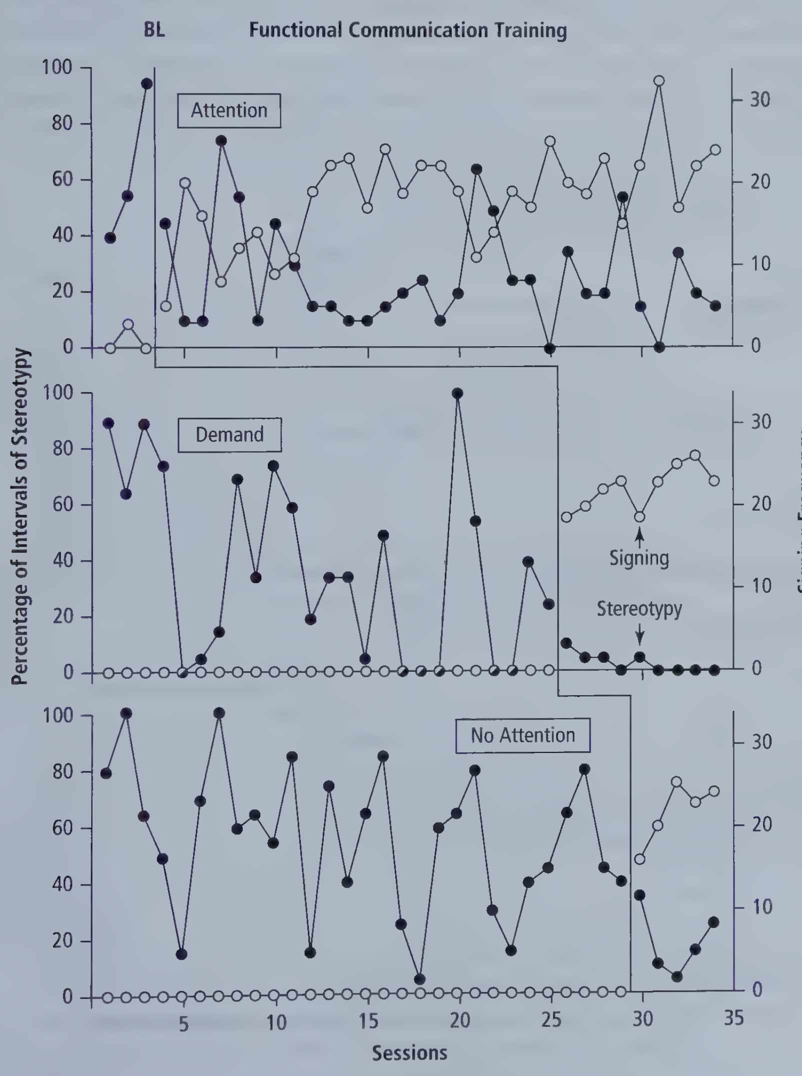


Figure 2. Occurrence of stereotypy for James across attention, demand, and no-attention conditions. Data are arrayed as the percentage of intervals of stereotypy on the left *y* axis and number of signs per sessions on the right *y* axis.

Figure 6.4 Graph with multiple data paths showing two different behaviors by one participant during baseline and training across three different conditions. Note the different dimensions and scaling of the dual vertical axes.

From "Analyzing the Multiple Functions of Stereotypical Behavior for Students with Autism: Implications for Assessment and Treatment," by C. H. Kennedy, K. A. Meyer, T. Knowles, and S. Shukla, 2000, *Journal of Applied Behavior Analysis, 33*, p. 565. Copyright 2000 by the Society for the Experimental Analysis of Behavior, Inc. Used by permission.

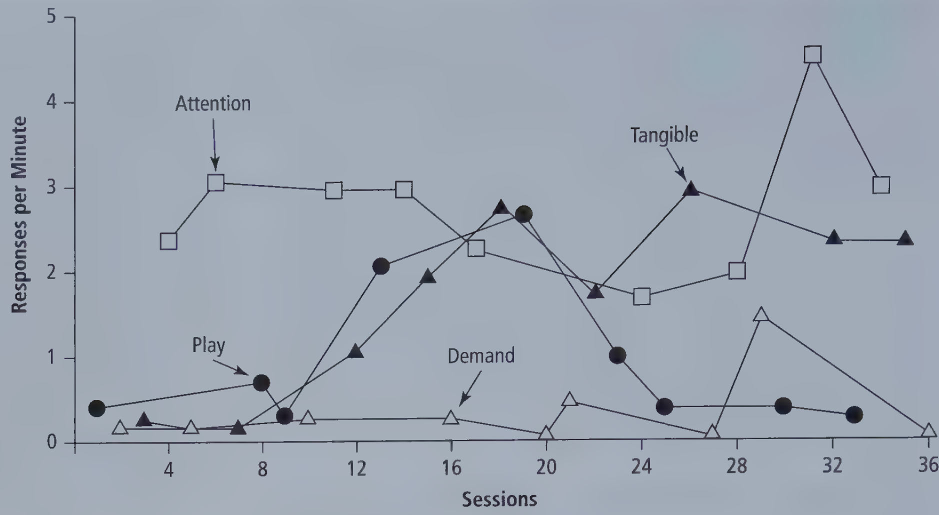


Figure 1. Rate of self-injurious behavior during the initial functional analysis.

Figure 6.5 Graph with multiple data paths showing the same behavior measured under four different conditions.

Based on "The Influence of Therapist Attention on Self-Injury during a Tangible Condition" by J. W. Moore, M. M. Mueller, M. Dubard, D. S. Roberts, and H. E. Sterling-Turner, 2002, *Journal of Applied Behavior Analysis, 35,* p. 285. Copyright 2002 by the Society for the Experimental Analysis of Behavior, Inc. Used by permission.

Figure 6.5 shows self-injurious responses per minute by a 6-year-old girl with developmental disabilities under four different conditions (Moore, Mueller, Dubard, Roberts, & Sterling-Turner, 2002). Graphing an individual's behavior under multiple conditions on the same set of axes allows direct visual comparisons of differences in absolute levels of responding at any given time as well as relative changes in performance over time.

Changing Values of an Independent Variable. Multiple data path graphs can also show changes in the target behavior (shown on one data path) relative to changing values of the independent variable (represented by a second data path). In each of the two graphs in Figure 6.6 one data path shows the duration of problem behavior (plotted against the left-hand *y*-axis scaled in seconds) relative to changes in noise level depicted by a second data path (plotted

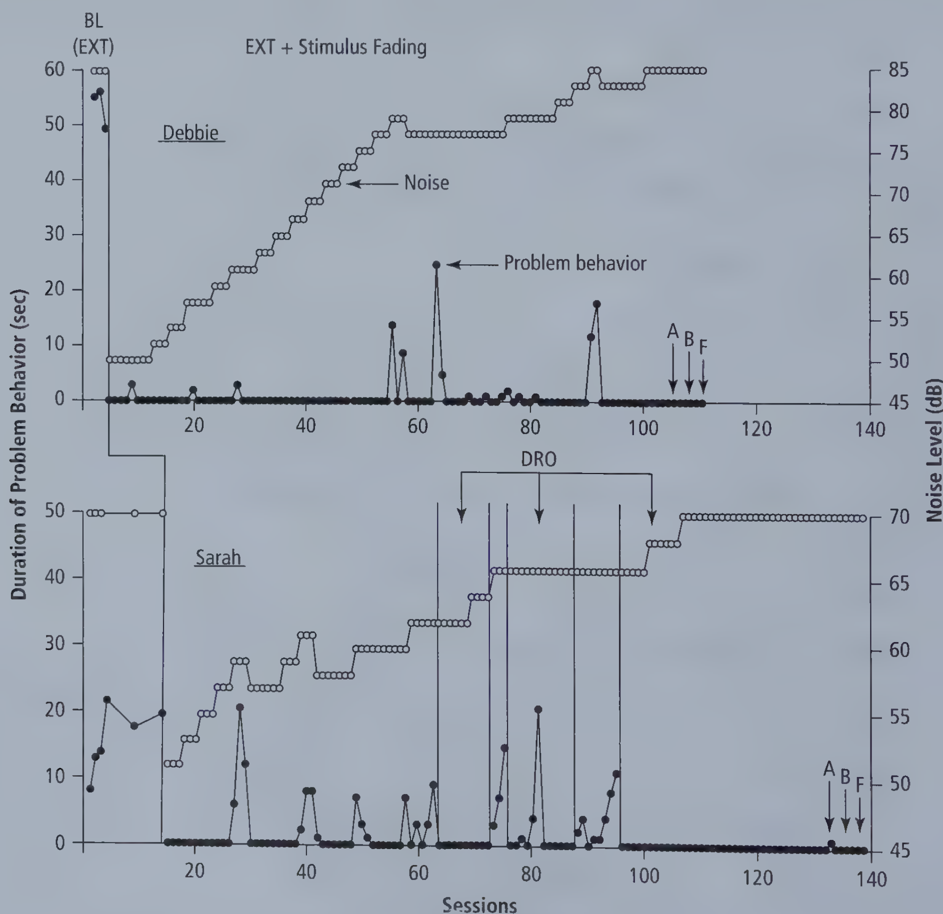


Figure 4. Results of Debbie's and Sarah's treatment evaluation. Sessions marked A and B near the end of treatment indicate two generalization probes in the natural environment; F indicates a follow-up probe.

Figure 6.6 Graph using two data paths to show the duration of problem behavior (dependent variable) by two adults with severe or profound intellectual disabilities as noise level was increased gradually (independent variable).

From "Functional Analysis and Treatment of Problem Behavior Evoked by Noise", B. E. McCord, B. A. Iwata, T. L. Galensky, S. A. Ellingson, and R. J. Thomson, 2001, Reproduced with permission of John Wiley & Sons Inc.

against the right-hand *y*-axis scaled in decibels) (McCord, Iwata, Galensky, Ellingson, & Thomson, 2001).

The Same Behavior of Two or More Participants. Multiple data paths are sometimes used to show the behavior of two or more participants on the same graph.

Depending on the levels and variability of the data encompassed by each data path, a maximum of four different data paths can be displayed effectively on one set of axes. However, there is no rule; Didden, Prinsen, and Sigafoos (2000) displayed five data paths in a single display. If too many data paths are displayed on the same graph, the benefits of making additional comparisons may be outweighed by the distraction of too much visual "noise." A combination of data paths and bars can help viewers distinguish measures of different behaviors or events. In a study evaluating a toilet training intervention, Greer, Neibert, and Dozier (2016) used one data path to represent the number self-initiations, another data path for the percentage of appropriate eliminations, and bars for the number of accidents (see Figure 6.7).

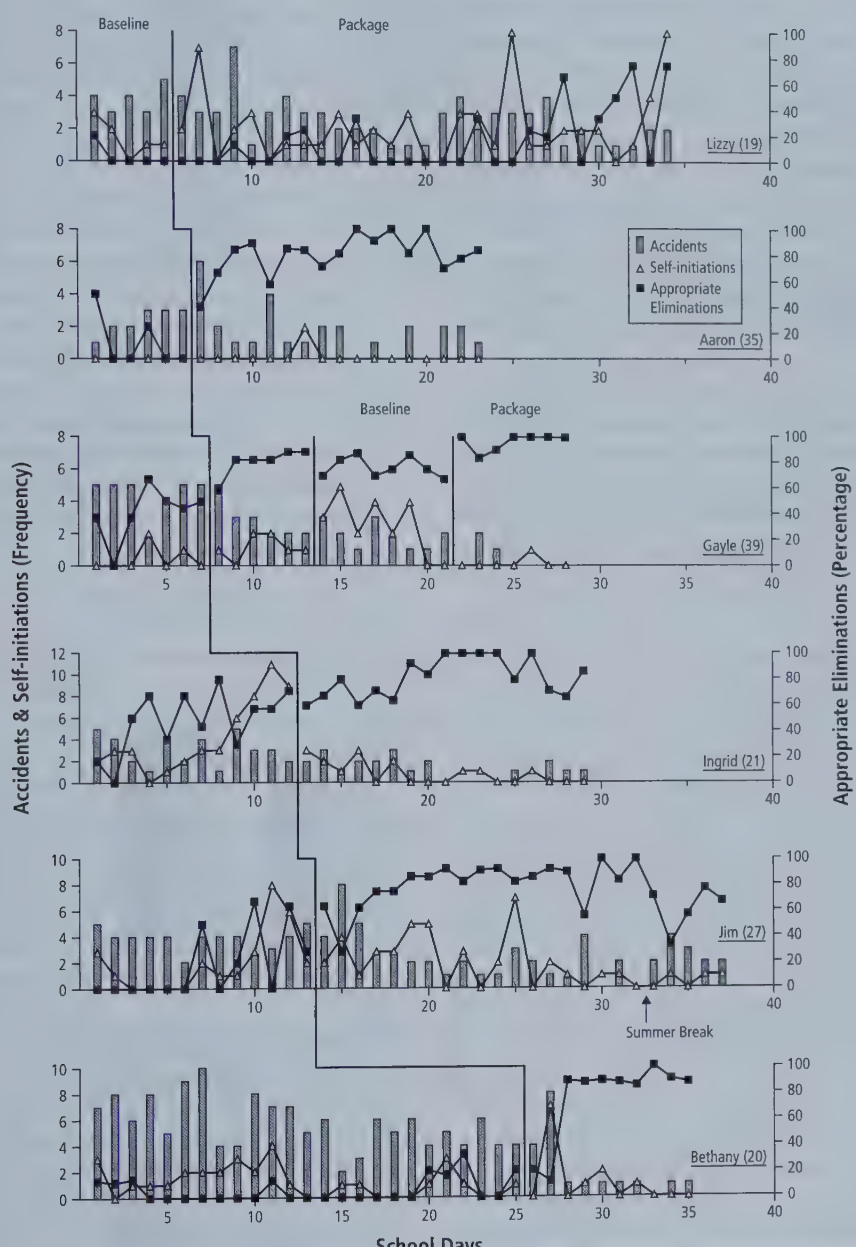


Figure 1. Results of the component analysis for children exposed to the toilet-training package after baseline. Each chid's age (in months) is noted parenthetically.

Figure 6.7 Graph using a combination of data path and bars to display changes in three response classes.

From "A Component Analysis of Toilet-Training Procedures Recommended for Young Children," by B. D. Greer, P. L. Neibert, and C. L. Dozier, 2016, *Journal of Applied Behavior Analysis, 49*, p. 76. Copyright 2016 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

Bar Graphs

The **bar graph**, or histogram, is a simple and adaptable format for graphically summarizing data. Bar graphs sacrifice the presentation of the variability and trends in data (which are apparent in a line graph) in exchange for the efficiency of summarizing and comparing large amounts of data in a simple, easy-to-interpret format.

Bar graphs are used frequently to summarize the performance of a participant or group of participants during different conditions. For example, Figure 6.8 shows the mean percentage of spelling worksheet items completed and completed correctly by four students during baseline and the combined generalization programming and maintenance conditions that followed training each child to recruit teacher attention while working (Craft, Alber, & Heward, 1998).

Bar graphs that present a measure of central tendency, such as the mean or median score for each condition, may mask important variability in the data. The range of measures represented by a mean or median can be shown with range bars (e.g., Putnam & Tiger, 2016). Figure 6.9 is an example of a bar graph showing both the range of measures and individual performances.

Bar graphs are also used to display discrete sets of data that are not related to one another by a common underlying dimension by which the horizontal axis could be scaled. For example, in a study analyzing the effects of establishing operations on preference assessments, Gottschalk, Libby, and Graff (2000) used bar graphs to show the percentage of trials in which four children reached toward and picked up different items (see Figure 6.10).

The *y*-axis on some bar graphs is scaled to negative values/below zero to show amount of behavior change above or below a baseline average or point in time (e.g., Greer, Neibert, & Dozier, 2016; Jessel, Hanley & Ghaemmaghami, 2016; Meindl, Ivy, Miller, Neef, & Williamson, 2013).

Cumulative Records

In the 1930s, Skinner invented the **cumulative recorder**, a device that automatically draws a graph of a subject's rate of behavior. In a book cataloging 6 years of experimental research on schedules of reinforcement, Ferster and Skinner (1957) described a cumulative record in the following manner:

> A graph showing the number of responses on the ordinate against time on the abscissa has proved to be the most convenient representation of the behavior observed in this research. Fortunately, such a "cumulative" record may be made directly at the time of the experiment. The record is raw data, but it also permits a direct inspection of rate and changes in rate not possible when the behavior is observed directly. . . . Each time the bird responds, the pen moves one step across the paper. At the same time, the paper feeds continuously. If the bird does not respond at all, a horizontal line is drawn in the direction of the paper feed. The faster the bird pecks, the steeper the line. (p. 23)

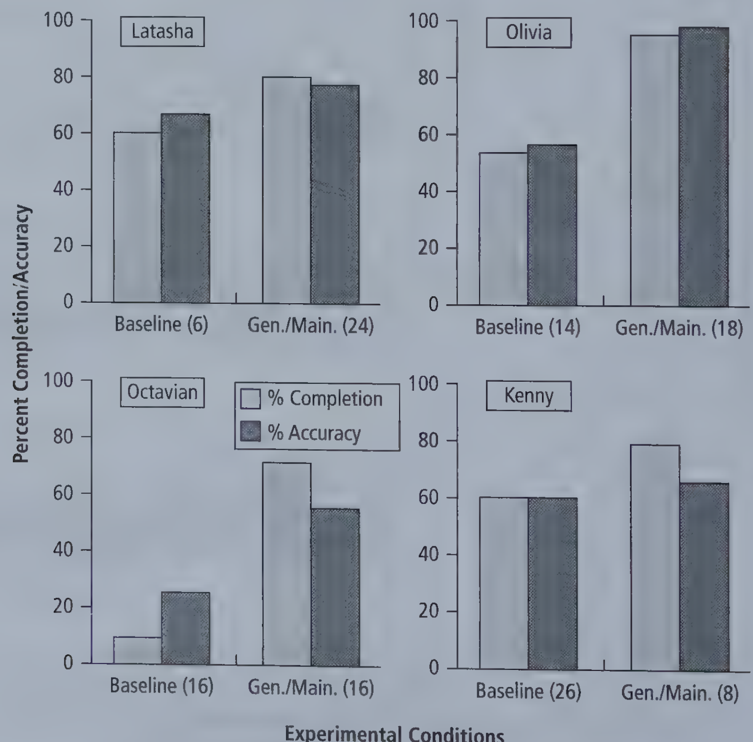


Figure 4. Mean percentage of spelling worksheet items completed and mean percentage of accuracy by each student during baseline and combined generalization programming and maintenance conditions. Numbers in parentheses show total number of sessions per condition.

Figure 6.8 Bar graph comparing mean levels for two dimensions of participants' performance between experimental conditions.

Based on "Teaching Elementary Students with Developmental Disabilities to Recruit Teacher Attention in a General Education Classroom: Effects on Teacher Praise and Academic Productivity" by M. A. Craft, S. R. Alber, and W. L. Heward, 1998, *Journal of Applied Behavior Analysis, 31*, p. 410. Copyright 1998 by the Society for the Experimental Analysis of Behavior, Inc.

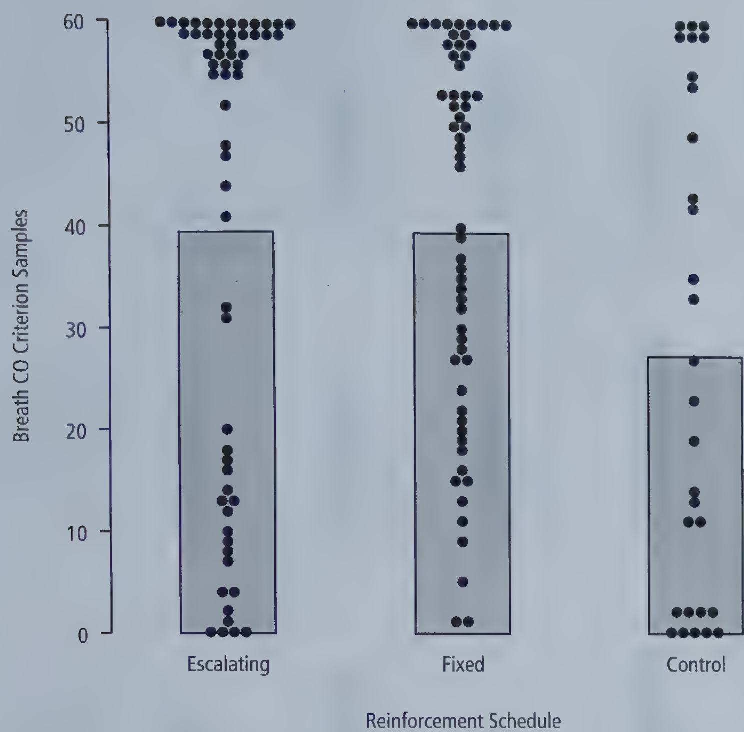


Figure 2. Total number of breath CO sample <3 ppm across each group. Each filled circle represents one participant. Gray area represents the mean percentage of breath CO samples <3 ppm for each group.

Figure 6.9 Bar graph comparing mean performance by three groups of participants under different experimental conditions, with data points representing each individual participant.

From "The Effects of Fixed Versus Escalating Reinforcement Schedules on Smoking Abstinence," by P. Romanowich and R. J. Lamb, 2015, *Journal of Applied Behavior Analysis, 48*, p. 33. Copyright 2015 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

For decades, cumulative records (graphs produced by a cumulative recorder) served as the primary form of data display in basic behavior analysis research laboratories around the world (Kangas & Cassidy, 2010). Although behavioral researchers seldom use cumulative recorders today, the cumulative graphs remain an important form of data display in both experimental and applied behavior analysis.[7]

A **cumulative record** is constructed by adding the number of responses recorded during each observation period (thus the term *cumulative*) to the total number of responses recorded during all previous observation periods. The *y*-axis value of any data point on a cumulative graph represents the total number of responses recorded since the beginning of data collection. Exceptions occur when the data path resets to the 0 value of the *y*-axis and begins its ascent again. Resets occur when the total number of responses has exceeded the upper limit of the *y*-axis scale, a new condition is implemented (Wilson & Gratz, 2016), or a performance criterion has been met (Williams, Perez-Gonzales, & Queiroz, 2005).

Figure 6.12 shows the cumulative number of spelling words mastered by a man with intellectual disabilities during baseline and two training conditions (Neef, Iwata, & Page, 1980). The individual mastered 1 word during the 12 sessions of baseline (social praise for correct spelling responses and rewriting incorrectly spelled words three times), 22 words under the interspersal condition (baseline procedures plus the presentation of a previously learned word after each unknown word), and 11 words under the high-density reinforcement condition (baseline procedures plus social praise given after each trial for task-related behaviors such as paying attention and writing neatly).

In addition to the total number of responses recorded at any given point in time, cumulative records show the overall and local response rates. Rate is the number of responses emitted per unit of time, usually reported as responses per minute in applied behavior analysis. An **overall response rate** is the average rate of response over a given time period, such as during a specific session, phase, or condition of an experiment. Overall rates are calculated by dividing the total number of responses recorded during the period by the number of observation periods indicated on the horizontal axis. In Figure 6.12 the overall response rates are 0.46 and 0.23 words mastered per session for the interspersal and high-density reinforcement conditions, respectively.[8]

On a cumulative record, the steeper the slope, the higher the response rate. To produce a visual representation of an overall rate on a cumulative graph, the first and last data points of a given series of observations should be connected with a straight line. A straight line connecting Points a and c in Figure 6.12 would represent the learner's overall rate of mastering spelling words during the interspersal condition. A straight line connecting Points a and e represents the overall rate during the high-density reinforcement condition. Relative rates of response can be determined by visually comparing one slope to another: the steeper the slope, the higher the rate of response. A visual comparison of Slopes a–c and a–e shows that the interspersal condition produced the higher overall response rate.

Response rates often fluctuate within a given period. The term **local response rate** refers to the rate of response during periods smaller than that for which an overall rate has been given. Over the last four sessions of the study shown in

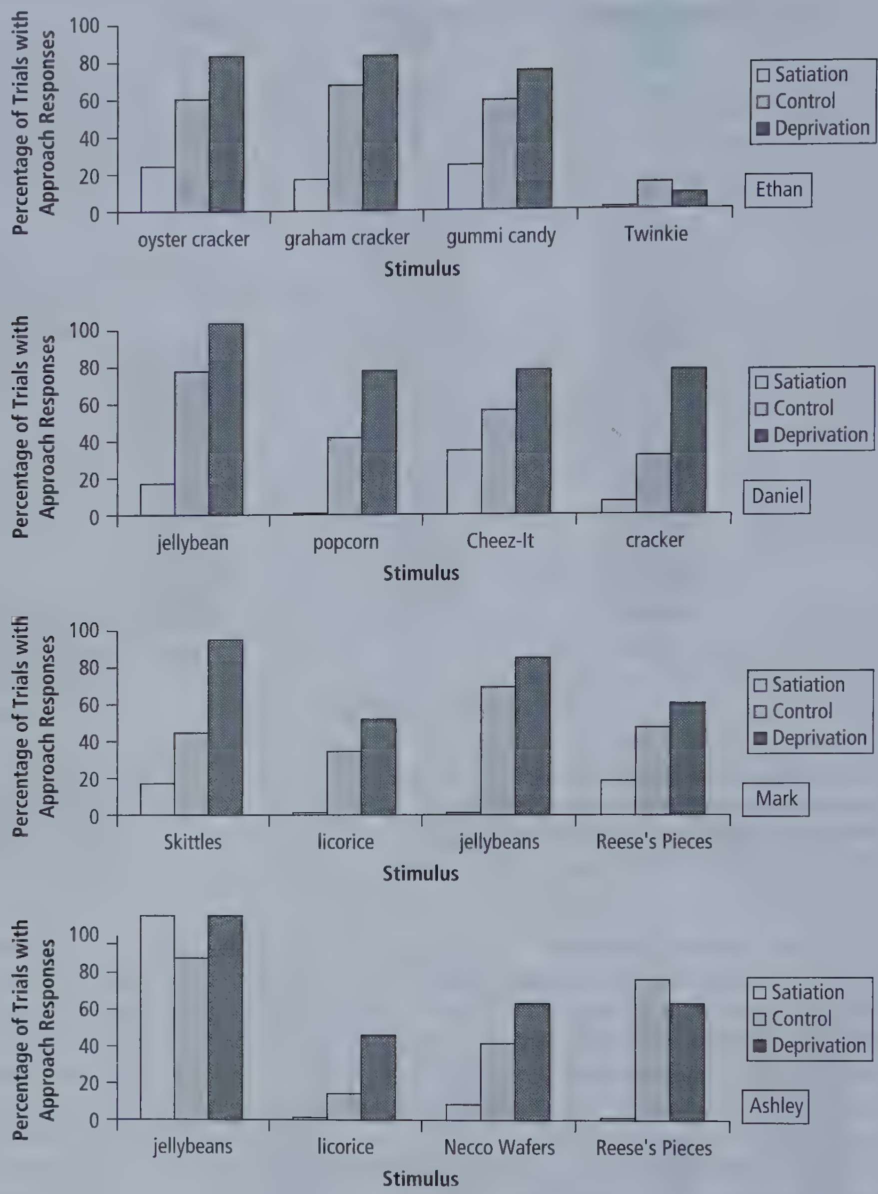


Figure 1. Percentage of approach responses across conditions for Ethan, Daniel, Mark, and Ashley.

Figure 6.10 Bar graph used to summarize, and display results of measurements taken under discrete conditions lacking an underlying dimension by which the horizontal axis could be scaled (e.g., time, duration of stimulus presentations).

Based on "The Effects of Establishing Operations on Preference Assessment Outcomes" by J. M. Gottschalk, M. E. Libby, and R. B Graff, 2000, *Journal of Applied Behavior Analysis, 33*, p. 87. Copyright 2000 by the Society for the Experimental Analysis of Behavior, Inc.

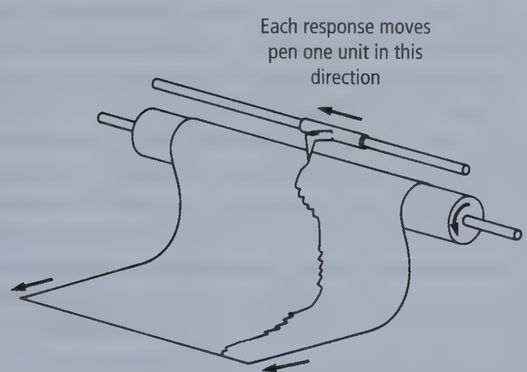


Figure 6.11 Diagram of a cumulative recorder.

Based on *Schedules of Reinforcement*, pp. 24–25, by C. B. Ferster and B. F. Skinner, 1957, Upper Saddle River, NJ: Prentice Hall. Copyright 1957 by Prentice Hall.

Figure 6.12, the learner exhibited a local rate of responding during interspersal training (Slope b–c) that was considerably higher than his overall rate for that condition. At the same time his performance during the final four sessions of the high-density reinforcement condition (Slope d–e) shows a lower local response rate than his overall rate for that condition.

A legend giving the slopes of some representative rates can aid considerably in the determination and comparison of relative response rates both within and across cumulative curves plotted on the same set of axes (e.g., see Kennedy & Souza, 1995, Figure 2). However, very high rates of responding are difficult to compare visually with one another on cumulative records.

> Although the rate of responding is directly proportional to the slope of the curve, at slopes above 80 degrees small differences in angle represent very large differences in

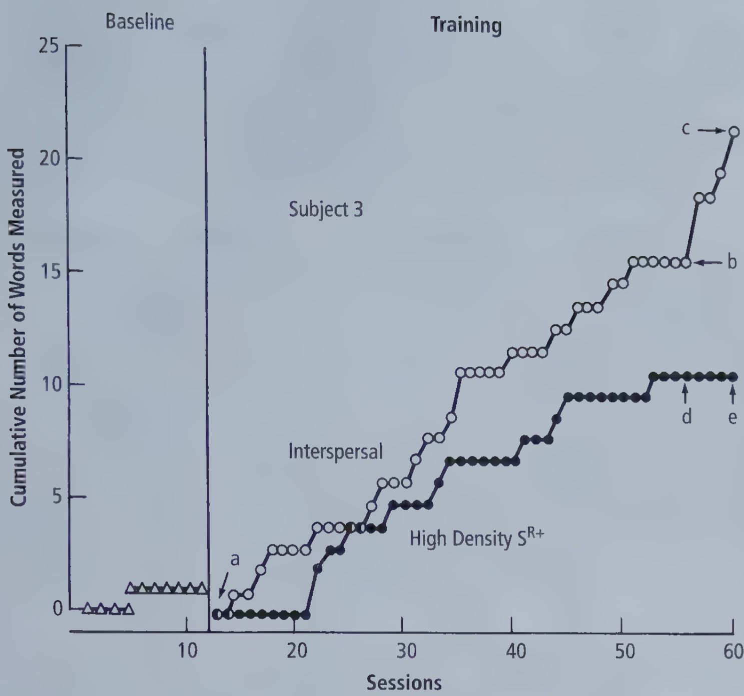


Figure 6.12 Cumulative graph of number of spelling words learned by a man with intellectual disabilities during baseline, interspersal training, and high-density reinforcement training. Points a–e have been added to illustrate the differences between overall and local response rates.

Based on "The Effects of Interspersal Training Versus High Density Reinforcement on Spelling Acquisition and Retention" by N. A. Neef, B. A. Iwata, and T. J. Page, 1980, *Journal of Applied Behavior Analysis, 13*, p. 156. Copyright 1980 by the Society for the Experimental Analysis of Behavior, Inc. Adapted by permission.

rate; and although these can be measured accurately, they cannot be evaluated easily by [visual] inspection. (Ferster & Skinner, 1957, pp. 24–25)

Even though cumulative records derived from continuous recording are the most directly descriptive displays of behavioral data available, two other features of behavior, in addition to the comparison of very high rates, can be difficult to determine on some cumulative graphs. One, although the total number of responses since data collection began can be easily seen on a cumulative graph, the number of responses recorded for any given session can be hard to ascertain, given the number of data points and the scaling of the vertical axis. Two, gradual changes in slope from one rate to another can be hard to detect on cumulative graphs.

A cumulative graph may be preferable to a noncumulative line graph in four situations. First, cumulative records are desirable when the total number of responses emitted over time is important or when progress toward a specific goal can be measured in cumulative units of behavior. The number of new words learned, dollars saved, or miles trained for an upcoming marathon are examples. One look at the most recent data point on the graph reveals the total amount of behavior up to that point in time.

Second, a cumulative graph may be more effective than a noncumulative graph when used as a source of feedback for clients. This is because both total progress and relative rate of performance are easily detected by visual inspection. Further, a cumulative graph does not reveal decelerating trends; it either ascends or plateaus. For individuals for whom a decelerating trend may signal failure, spark embarrassment, incite retaliation by members of a peer group, or otherwise affect them negatively, a cumulative graph, with its ascending trend (or plateaus), may mitigate concerns.

Third, a cumulative record should be used when the target behavior can occur or not occur just once per observation session. In these instances, the effects of an intervention may be easier to detect on a cumulative graph. Figure 6.13 shows the same data plotted on a noncumulative graph and a cumulative graph. The cumulative graph clearly shows a relation between behavior and intervention, whereas the noncumulative graph gives the impression of greater variability in the data than really exists.

Fourth, cumulative records can "reveal the intricate relations between behavior and environmental variables" (Johnston & Pennypacker, 2009, p. 317). Figure 6.14 is an excellent example of how a cumulative graph enables a detailed analysis of behavior change (Hanley, Iwata, & Thompson, 2001). By plotting the data from single sessions cumulatively by 10-second intervals, the researchers revealed patterns of responding not shown by a graph of session-by-session data. Comparing the data paths for the three sessions for which the results were graphed cumulatively (Mult #106, Mixed #107, and Mixed #112) revealed two undesirable patterns of responding (in this case, pushing a switch that operated a voice output device that said "talk to me, please" as an alternative to self-injurious behavior and aggression) that are likely to occur during mixed schedules, and the benefits of including schedule-correlated stimuli (Mult #106).

Ratio Charts

All of the line graphs discussed so far have been equal-interval graphs on which the distance between any two consecutive points on each axis is always the same. On the horizontal axis, the distance between data on Day 1 and Day 2 has the same distance as Day 11 to Day 12; on the vertical axis, the distance between 10 and 20 responses per minute visually matches the distance between 35 and 45 responses per minute. On an equal-interval graph, equal vertical distances on the vertical axis represent equal amounts of behavior change, whether the performance increases or decreases.

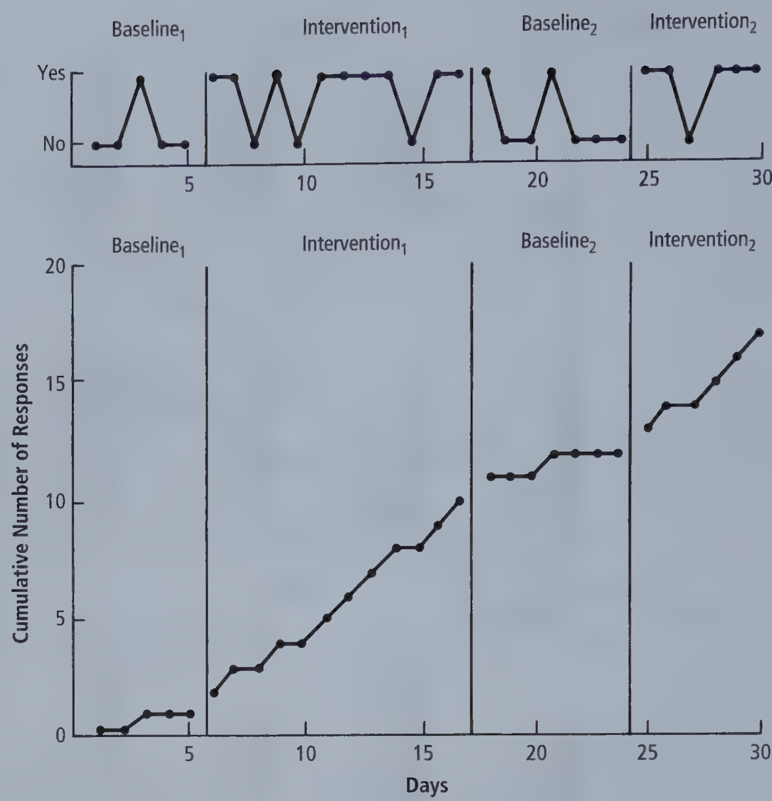


Figure 6.13 Same set of hypothetical data plotted on noncumulative and cumulative graphs. Cumulative graphs more clearly reveal patterns of and changes in responding for behaviors that can occur only once during each period of measurement.

Based on *Working with Parents of Handicapped Children*, p. 100, by W. L. Heward, J. C. Dardig, and A. Rossett, 1979, Columbus, OH: Charles E. Merrill. Copyright 1979 by Charles E. Merrill.

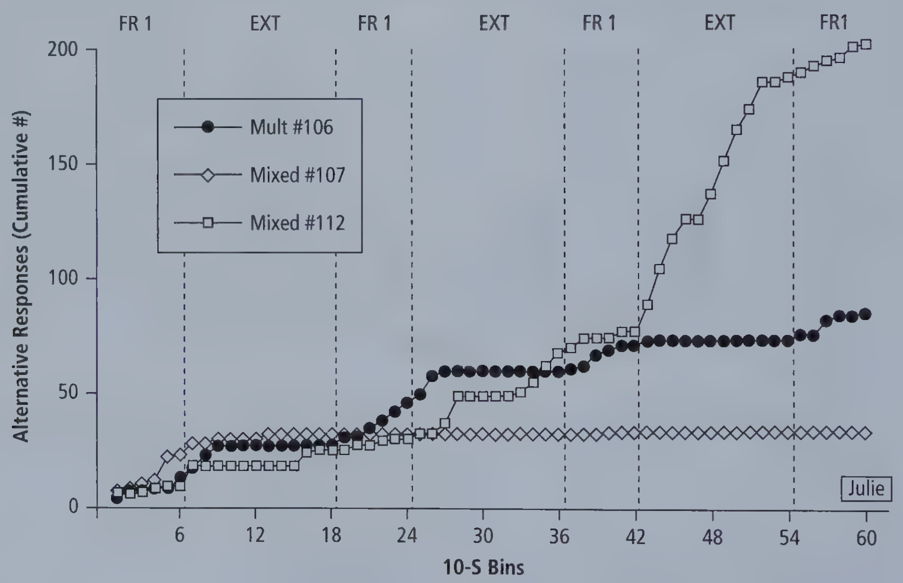


Figure 4. Cumulative number of alternative responses across schedule components for three sessions from Julie's assessment of schedule-correlated stimuli. Open symbols represent two mixed-schedule sessions; filled symbols represent a single multiple-schedule session.

Figure 6.14 Cumulative record used to make a detailed analysis and comparison of behavior across components of multiple- and mixed-reinforcement schedules within specific sessions of a study.

Based on "Reinforcement Schedule Thinning Following Treatment with Functional Communication Training" by G. P. Hanley, B. A. Iwata, and R. H. Thompson, 2001, *Journal of Applied Behavior Analysis, 34*, p. 33. Copyright 2001 by the Society for the Experimental Analysis of Behavior, Inc. Used by permission.

Another way to examine behavior change is by proportional or relative change. Ratio scales are well suited for displaying proportional change. On a **ratio scale,** equal vertical distances correspond to equal ratios of change in the variable being measured. Because behavior is measured and charted over time, which progresses in equal intervals, the horizontal axis is marked off in equal intervals and only the vertical axis has a ratio scaling. Hence, the term *ratio chart* refers to graphs where one axis is scaled proportionally.

Ratio charts offer a different perspective, namely, detecting relative behavior change as opposed to absolute change. For example, on a ratio chart a doubling of response rate from 4 to 8 per minute would appear the same as a doubling of 50 to 100 responses per minute. Likewise, a decrease in responding from 75 to 50 responses per minute (a decrease of one third) would occupy the same distance on the vertical axis as a change from 12 to 8 responses per minute (a decrease of one third).

Figure 6.15 shows the same data plotted on an equal-interval graph (sometimes called *linear, arithmetic,* or *add-subtract charts*) and on a ratio chart (sometimes called a *semilogarithmic chart*). The behavior change that appears as an

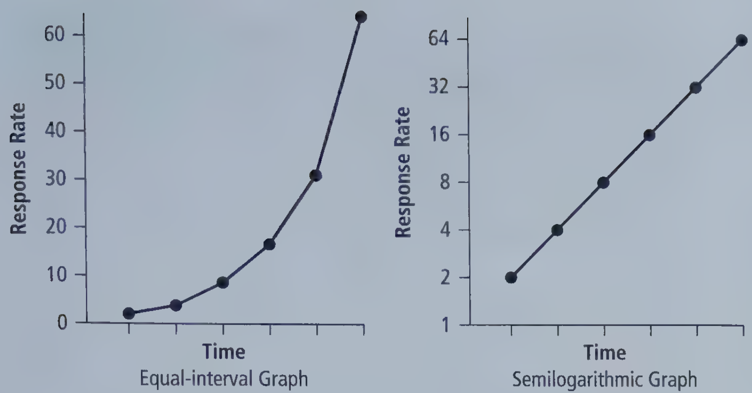


Figure 6.15 Same set of data plotted on equal-interval arithmetic scale (left) and on equal-proportion ratio scale (right).

exponential curve on the arithmetic chart is a straight line on the ratio chart. The vertical axis in the ratio chart in Figure 6.15 is scaled by log-base-2 or ×2 cycles, which means each cycle going up the vertical axis represents a times-2 increase (i.e., a doubling or 100% increase) of the cycle below it.

Standard Celeration Chart

Ogden Lindsley developed the **Standard Celeration Chart** to provide a standardized means of charting and analyzing how frequency of behavior changes over time (Lindsley, 1971). The Standard Celeration Chart is a standard ratio chart with six ×10 cycles on the vertical axis that can accommodate response rates as low as 1 per 24 hours (0.000695 per minute) or as high as 1000 per minute. There are four standard charts, differentiated from one another by the scaling on the horizontal axis: a daily chart with 140 calendar days, a weekly chart, a monthly chart, and a yearly chart. The daily chart shown in Figure 6.16 is used most often.

The size of the chart and the consistent scaling of the vertical axis and horizontal axis do not make the Standard Celeration Chart *standard*, as is commonly believed. What makes the Standard Celeration Chart standard is its consistent

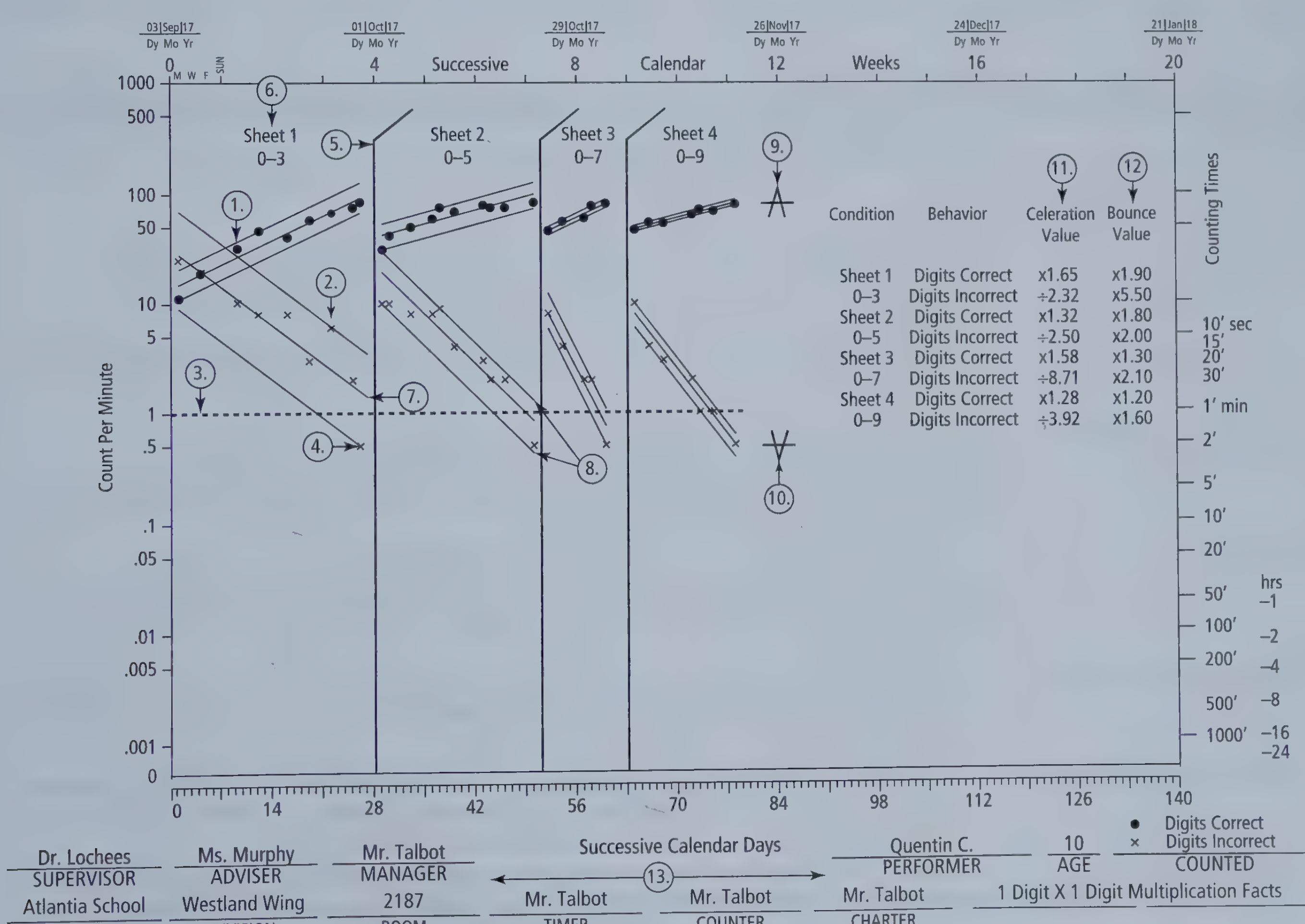


Figure 6.16 Standard Celeration Chart showing basic charting conventions. (See Table 6.1 for descriptions of numbered features.)

Model Standard Celeration Chart created by Douglas E. Kostewicz, Ph.D., BCBA-D, University of Pittsburgh. Used with permission.

display of *celeration*, a linear measure of frequency change across time, a factor by which frequency multiplies or divides per unit of time. The terms *acceleration* and *deceleration* are used to describe accelerating performances or decelerating performances.

A line drawn from the bottom left corner to the top right corner has a slope of 34° on all Standard Celeration Charts. The slope has a celeration value of ×2 (read as "times-2"). Celerations are expressed with a multiplication or division sign. A ×2 celeration represents a doubling in frequency per celeration period. The celeration period for the daily chart is per week; per month for the weekly chart, per 6 months for the monthly chart, and per 5 years for the yearly chart.

An instructional decision-making system, called **precision teaching**, has been developed for use with the Standard Celeration Chart. Precision teaching is predicated on the position that (a) learning is best measured as a change in response rate, (b) learning most often occurs through proportional changes in behavior, and (c) past changes in performance can project future learning. Figure 6.16 shows how precision teaching could be implemented to help a student working on 1-digit by 1-digit multiplication facts. Several times per week the student is given a worksheet of 1-digit by 1-digit multiplication problems and told to write the answer to as many problems as he can in 1 minute. Practice sheets during the initial phase consist of problems with digits 0 through 3 (e.g., 3×3, 2×1, 0×1). When the student attains the dual aims of 80 correct digits and 0 incorrect digits per minute, he starts practicing problems with increased range of digits and his performance is assessed with practice sheets sampling more diverse populations of problems. He completes the skill by reaching the aim on the final set of worksheets sampling all possible 1-digit by 1-digit multiplication problems. The Standard Celeration Chart shows the student took the longest to reach aim on the first set of sheets but attained each successive aim more quickly and with fewer errors. The reduced variability, called *bounce* by precision teachers, for corrects for each successive type of practice sheet suggests tighter stimulus control. Table 6.1 provides names and descriptions of chart features in Figure 6.16.[9]

Scatterplots

A **scatterplot** is a graphic display that shows the relative distribution of individual measures in a data set with respect to the variables depicted by the *x*- and *y*-axes. Data points on a scatterplot are unconnected. Scatterplots show how much changes in the value of the variable depicted by one axis correlate with changes in the value of the variable represented by the other axis. Patterns of data points falling along lines on the plane or clusters suggest certain relationships. For example, a scatterplot

TABLE 6.1 Features of the Standard Celeration Chart as Used by Precision Teaching (numbers refer to Figure 6.16)

Chart Feature	Description and Charting Convention
1. Acceleration target frequency	Responses to accelerate noted by a solid dot on the appropriate day line.
2. Deceleration target frequency	Responses to decelerate noted by an X on the appropriate day line.
3. Time bar	Amount of recording time per day noted by a solid horizontal line on the charted day.
4. No count (zero)	No behaviors intended to accelerate (i.e., dots) or decelerate (i.e., X) occurred during the recording period and charted at a distance ÷2 below the time bar.
5. Condition change line	Vertical line drawn between the last charted day of one intervention and the first charted day of a new intervention.
6. Condition change label	Brief description of condition/intervention.
7. Celeration lines	A straight line representing celeration or deceleration of five or more charted days with each condition. Drawn by free hand on paper charts; automatically calculated with linear regression algorithm and plotted on some digital charts.
8. Bounce lines	A measure of performance variability. Parallel straight lines above and below a celeration line that encompass the majority of performance. The up bounce line is placed through the upmost point (excluding outliers); the down bounce line is placed through the downmost point (excluding outliers).
9. Acceleration frequency aim	A symbol representing the desired frequency aim and date to achieve it for the behavior targeted for acceleration. Drawn with horizontal bar on the desired frequency and date and intersecting caret pointing up.
10. Deceleration frequency aim	A symbol representing the desired frequency aim and date to achieve it for the behavior targeted for deceleration. Drawn with horizontal bar on the desired frequency and date and intersecting caret pointing down.
11. Celeration value	A measure of performance change per minute per week. Accelerations (X) and decelerations (÷) represented on the chart or text show the change in performance.
12. Bounce values	The total amount of bounce (i.e., variability) around celeration and deceleration in a condition. Always written as an "X" value.

TABLE 6.1 (continued)

Chart Feature	Description and Charting Convention	
13. Charted blanks	Supervisor	Sees the chart monthly, advises the adviser or manager
	Adviser	Advises the manager or performer
	Manager	Works with the performer daily
	Organization	Where the recorded performance takes place
	Division	Location or subdivision within the organization
	Room	Room where the recording takes place
	Counter	Counts the performer's behavior
	Performer	Learner whose performance appears on the chart
	Charter	Person who creates chart with recorded performance data
	Counted	Description of performer's counted behavior
	Small Unlabeled Blanks	Optional: performer's age
	Large Unlabeled Blanks	Optional: additional relevant information

Douglas Kostewicz, Ph.D., BCBA, University of Pittsburgh. Used with permission.

graph presented by Meindl et al. (2013) suggested that the more fluent a student's performance on a set of training flash cards, the greater the decrease in fluency with a generalization set of cards (see Figure 6.17). This finding supported the researchers' hypothesis that the student's responses to the training set flash-cards were controlled in part by irrelevant stimulus features of the cards (e.g., specific words in specific locations on the cards), which inhibited generalization to new cards.

Scatterplots can also reveal relationships among different subsets of data. For example, Boyce and Geller (2001) created the scatterplot shown in Figure 6.18 to see how the behavior of individuals from different demographic groups related to a ratio of driving speed and following distance that represents one element of safe driving (e.g., the proportion of data points for young males falling in the at-risk area of the graph compared to the proportion of drivers from other groups). Each data point

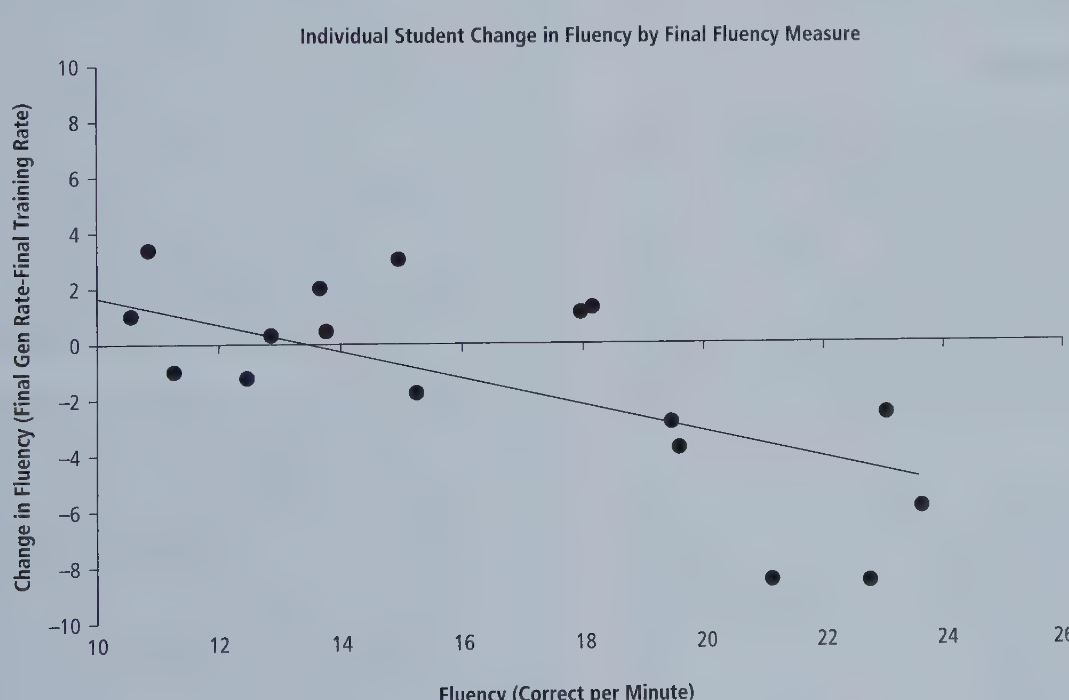


Figure. 2 Scatterplot depicting the relation between student fluency on the final training timing and the change in fluency from the final timing to the generalization test

Figure 6.17 Scatterplot depicting the relation between student fluency on the final training timing and the change in fluency from the final timing to a generalization test.

From "An Examination of Stimulus Control in Fluency-based Strategies: SAFMEDS and Generalization," by J. N. Meindl, J. W. Ivy, N. Miller, N. A. Neef, and R. L. Williamson, 2013, *Journal of Behavioral Education, 22,* p. 238. Copyright 2013 by Springer Sciences+Business Media. Used by permission.

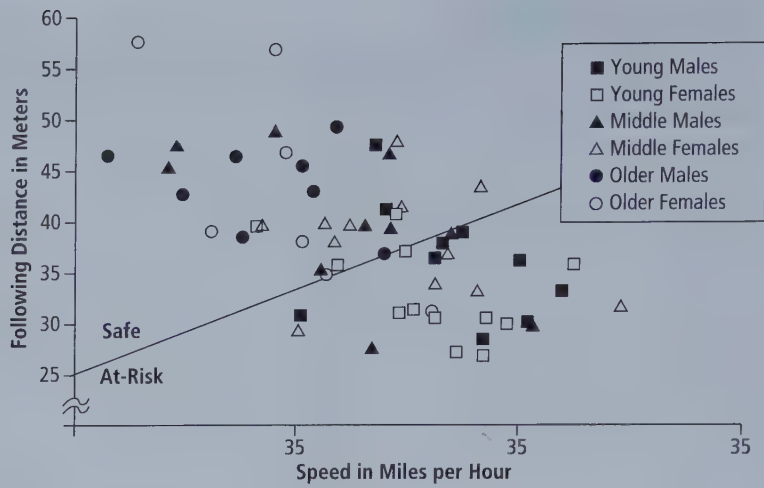


Figure 6.18 Scatterplot showing how the behavior of individuals from different demographic groups relates to a standard measure of safe driving.

Based on "A Technology to Measure Multiple Driving Behaviors without Self-Report or Participant Reactivity" by T. E. Boyce and E. S. Geller, 2001, *Journal of Applied Behavior Analysis, 34*, p. 49. Copyright 2001 by the Society for the Experimental Analysis of Behavior, Inc. Used by permission.

shows a single driver's behavior in terms of speed and following distance and whether the speed and following distance combination is considered safe or at risk for accidents. Such data could be used to target interventions for certain demographic groups.

Applied behavior analysts sometimes use scatterplots to discover the temporal distribution of a target behavior (e.g., Kahng et al., 1998; Symons, McDonald, & Wehby, 1998; Touchette, MacDonald, & Langer, 1985). Touchette and colleagues described a procedure for observing and recording behavior that produces a scatterplot that graphically shows whether the behavior's occurrence is typically associated with certain time periods. The use of scatterplot recording is described further in Chapter 27.

Other Graphic Displays

Applied behavior analysts have used a wide variety of other graphic formats to display and communicate behavioral data. For example, Raiff, Jarvis, and Dallery (2016) created a graphic display consisting of a horizontal sequence of boxes in different shades to illustrate patients' adherence to antidiabetic medication routines during baseline and treatment conditions. Romanowich and Lamb (2015) illustrated and compared participants' ratings on a treatment acceptability scale with a horizontal bar graph. Still other graphic formats are *bubble graphs* (e.g., Schultz, Kohn, & Musto, 2017), *dot charts* (see Figure A in Box 6.1), and *frequency profiles* (Epstein, Mejia, & Robertson, 2017).

Applied behavior analysis researchers have devised creative combinations of graphic devices to present data in ways that invite insights, understanding, and analyses of their findings. Interesting and effective examples can be found in Cammilleri and Hanley (2005 [Figure 29.10 in this text]; Fahmie, Iwata, and Mead (2016, Figure 1); Ingvarsson, Kramer, Carp, Pétursdóttir, and Macias (2016, Figures 1 and 2); Jessel et al. (2016, Figure 5); and Jones, Carr, and Feeley (2006, Figure 10).[10]

CONSTRUCTING LINE GRAPHS

The skills required for constructing effective, distortion-free graphic displays are as important as any in the behavior analyst's repertoire.[11] As applied behavior analysis has developed, so have certain stylistic conventions and expectations regarding the construction of graphs. An effective graph presents the data accurately, completely, and clearly, and makes the viewer's task of understanding the data as easy as possible. The graph maker must strive to fulfill each of these requirements while remaining alert to features in the graph's design or construction that might create distortion and bias—either the graph maker's or that of a future viewer when interpreting the extent and nature of the behavior change depicted by the graph.

Despite the graph's prominent role in applied behavior analysis, relatively few detailed treatments of how to construct behavioral graphs have been published. Notable exceptions have been chapters by Parsonson and Baer (1978, 1986) and by Spriggs, Lane, and Gast (2014); discussion of graphic display tactics by Johnston and Pennypacker (2009); and a recent critique of graphs published in behavior analytic journals by Kubina, Kostewicz, Brennan, and King (2015). Recommendations from these sources and others (American Psychological Association, 2009; Tufte, 2001) contributed to the preparation of this section. Additionally, hundreds of graphs published in applied behavior analysis journals were examined in an effort to discover those features that most clearly communicate the story told by the data.

Although there are few hard-and-fast rules for constructing graphs, adhering to the following conventions will result in clear, well-designed graphic displays consistent in format and appearance with current practice. Although most of the recommendations are illustrated by graphs presented throughout this text, we created Figures 6.19 and 6.20 to serve as models for most of the practices suggested here. The recommendations generally apply to all behavioral graphs. However, each data set and the conditions under which the data were obtained present their own challenges to the graph maker.

Drawing, Scaling, and Labeling Axes

Ratio of the Vertical and Horizontal Axes

The relative length of the vertical axis to the horizontal axis, in combination with the scaling of the axes, determines the degree

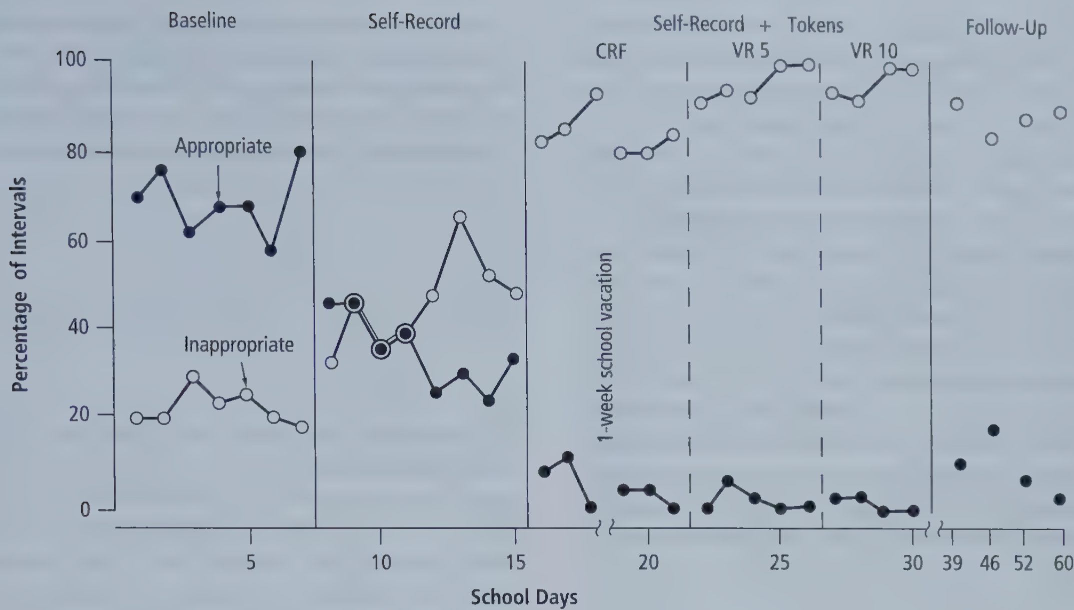


Figure 1. Percent of 10-second intervals of appropriate and inappropriate behavior by an 8-year-old boy. Each interval was scored as appropriate, inappropriate, or neither; thus the two behaviors do not always total 100%.

Figure 6.19 Graph of hypothetical data illustrating various conventions and guidelines for constructing behavioral graphs.

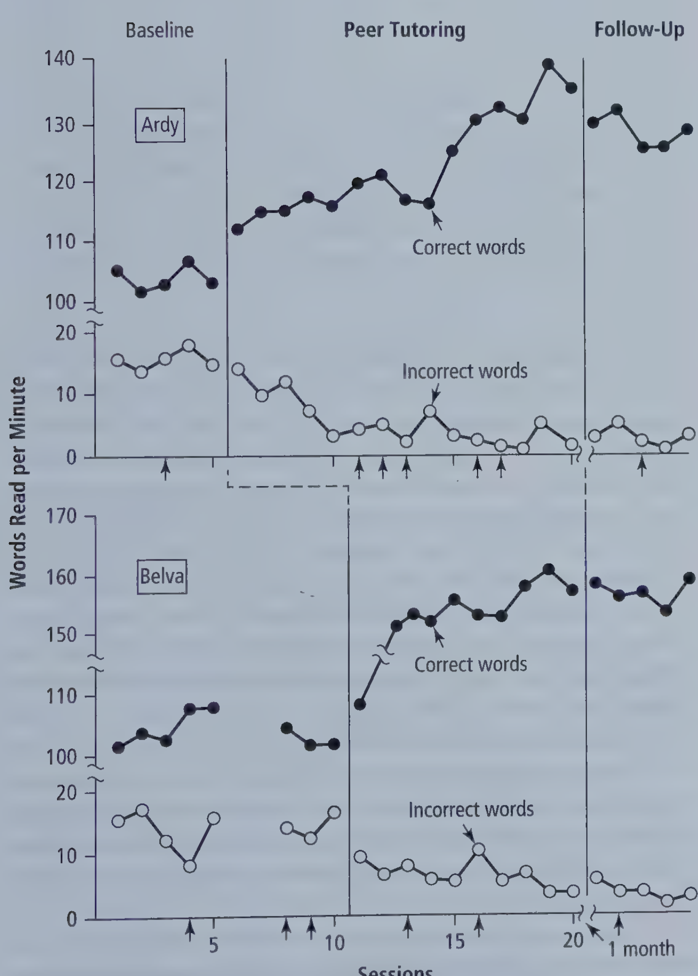


Figure 1. Number of words read correctly and incorrectly by Ardy and Belva during 1-min probes following each session. Arrows below horizontal axes indicate sessions in which student used reading material brought from home.

Figure 6.20 Graph of hypothetical data illustrating various conventions and guidelines for constructing behavioral graphs.

to which a graph will accentuate or minimize variability in a given data set. The legibility of a graph is enhanced by a balanced ratio between the height and width so that data points are neither too close together nor too spread apart. Recommendations in the behavioral literature for the ratio, or relative length, of the vertical axis to the horizontal axis range from 5:8 (Johnston & Pennypacker, 2009) to 3:4 (Katzenberg, 1975). Tufte (2001), whose book *The Visual Display of Quantitative Information* is a wonderful storehouse of guidelines and examples of effective graphing techniques, also recommends a vertical axis–to–horizontal axis ratio of 5:8.

A vertical axis approximately two-thirds the length of the horizontal axis works well for most behavioral graphs. When multiple sets of axes will be presented one atop another in a single figure and/or when the number of data points to be plotted on the horizontal axis is very large, the length of the vertical axis relative to the horizontal axis can be reduced (as shown in Figures 6.2 and 6.7).

Scaling the Horizontal Axis

The horizontal axis should be marked in equal intervals, with each unit representing from left to right the chronological succession of equal time periods or response opportunities in which the behavior was (or will be) measured and from which an interpretation of behavior change is to be made (e.g., days, sessions, trials). It is not necessary to mark each point on the *x*-axis. Avoid visual clutter by placing *tick marks* at regularly spaced increments on the outside of the axis and numbering them accordingly (e.g., by 5s, 10s, 20s).

When two or more sets of axes are stacked vertically, and each horizontal axis represents the same time frame, unnumbered tick marks on the horizontal axes of the upper tiers are permissible. However, each horizontal axis should have the tick marks corresponding to those on the bottom tier to facilitate comparison of performance across tiers at any given point in time (see Figure 6.4).

Representing Discontinuities of Time on the Horizontal Axis

Behavior change, its measurement, and all manipulations of treatment or experimental variables occur within and across time. Therefore, time is a fundamental variable in all experiments that should not be distorted or arbitrarily represented in a graphic display. Each equally spaced unit on the horizontal axis should represent an equal passage of time. Discontinuities in the progression of time should be indicated on the horizontal axis by a *scale break*: an open, unconnected portion of the axis with a squiggly line at each end. Scale breaks on the *x*-axis can also be used to signal periods when data were not collected or when otherwise regularly spaced data points represent consecutive measurements made at unequal intervals (see the numbering of school days for the follow-up condition in Figure 6.19).

When measurement occurs across consecutive observations of opportunities for the target behavior to occur (e.g., stories read, meals, tennis serves) rather than standard units of time, the horizontal axis still serves as a visual representation of the progression of time because the data plotted above it are recorded one after the other. The text accompanying such a figure should indicate the real time in which the consecutive measurements were made (e.g., "Two or three peer tutoring sessions were conducted each school week"), and any irregularities or discontinuities in that time context should be clearly indicated or marked with scale breaks (see Figure 6.20).

Labeling the Horizontal Axis

The dimension by which the horizontal axis is scaled is identified in a brief label centered below and parallel to the axis.

Scaling the Vertical Axis

The most significant feature of an equal-interval graph's representation of behavior change (or lack thereof) is its vertical axis scale. Common practice is to mark the origin at 0 (on cumulative graphs the bottom on the vertical axis must be 0) and then to mark off the vertical axis to accommodate all values represented in the data set and the range of socially important/desired performance. Increasing the distance on the vertical axis between each unit of measurement magnifies the variability in the data, whereas contracting the units of measurement on the vertical axis minimizes the portrayal of variability in the data set. The graph maker should plot the data set against several different vertical axis scales, watching for distortion of the graphic display that might lead to inappropriate interpretations (Dart & Radley, 2017).

When zero values make no sense, truncate the *y*-axis. "Charting is about being true to the data. Some data never fall to zero—the body temperature of a living person, for instance" (Yanofsky 2015). Edward Tufte (2017) agrees: "Don't spend a lot of empty vertical space trying to reach down to the zero point at the cost of hiding what is going on in the data line itself" (n.p.).

The social significance of various levels of behavior change for the behavior being graphed should be considered when scaling the vertical axis. If relatively small numerical changes in performance are socially significant, the *y*-axis scale should reflect a smaller range of values. For example, to display data most effectively from a training program in which an industrial employee's percentage of correctly executed steps in a safety checklist increased from an unsafe preintervention range of 85% to 90% to a nearly accident-free postintervention level of 98% to 100%, the vertical axis should focus on the 80% to 100% range. However, the scaling of the vertical axis should be contracted when small numerical changes in behavior are not socially important and the degree of variability obscured by the compressed scale is of little interest.

Horizontal numbering of regularly spaced tick marks on the outside of the vertical axis reduces visual clutter and facilitates reading the scale. The vertical axis should not be extended beyond the tick mark indicating the highest value on the axis scale.

When the data set includes several measures of 0, start the vertical axis at a point slightly above the horizontal axis to keep data points from falling directly on the axis. This produces a neater graph and helps the viewer discriminate 0-value data points from those representing measures close to 0 (see Figure 6.19).

In most instances, scale breaks should not be used on the vertical axis. However, when two sets of data with widely different and nonoverlapping ranges are displayed against the same *y*-axis, a scale break can be used to separate the range of measures encompassed by each data set (see Figure 6.20).

In multiple-tier graphs, equal distances on each vertical axis should represent equal changes in behavior to aid the comparison of data across tiers. Also, whenever possible, similar positions on each vertical axis of multiple-tier graphs should represent similar absolute values of the dependent variable. When the differences in behavioral measures from one tier to another would result in one or more overly long vertical axes, scale breaks can be used to highlight the difference in absolute values and aid point-to-point comparisons across tiers.

When the range values to be plotted on the vertical axis exceed what could be displayed on an axis of proportional length to the horizontal axis, it may be possible to split the range across two or more tired graphs. For example, in a study on compliance, Cook, Rapp, and Schulze (2015) presented their results in a three-tiered graph. (see Figure 9.11 this text). The vertical axis of the top tier is scaled in seconds, the middle tier in minutes, and the bottom tier in hours. The horizontal axis of each descending tier represented a continuation of time as measured by experimental trials.

Labeling the Vertical Axis

A brief label, printed and centered to the left and parallel to the vertical axis, should identify the dimension by which the axis is scaled. On multiple-tiered graphs, one label identifying the dimension portrayed on all of the vertical axes can be centered along the axes as a group. Additional labels identifying the different behaviors (or some other relevant aspect) graphed within each set of axes are sometimes printed to the left and parallel to each vertical axis. These individual tier labels should be printed to the right of and in smaller sized font than the label identifying the dimension by which all of the vertical axes are scaled.

Identify Experimental Conditions

Condition Change Lines

Vertical lines extending upward from the horizontal axis indicate changes in treatment or experimental procedures. Condition change lines should be placed after (to the right of) the data point representing the last measure prior to the change in conditions signified by the line and before (to the left of) the data point representing the first measure obtained after the change in procedure. In this way, data points fall clearly on either side of change lines and never on the lines themselves. Drawing condition change lines to a height equal to the height of the vertical axis helps the viewer estimate the value of data points near the top of the vertical axis range.

Condition change lines can be drawn with either solid or dashed lines. However, when an experiment or a treatment program includes relatively minor changes within an ongoing condition, a combination of solid and dashed lines should be used to distinguish the major and minor changes in conditions. For example, the solid lines in Figure 6.19 change from baseline, to Self-Record, to Self-Record+Tokens, to Follow-Up conditions, and dashed lines indicate changes in the schedule of reinforcement from CRF, to VR 5, to VR 10 within the Self-Record+Tokens condition (CRF and VR schedules of reinforcement are described in Chapter 13).

When the same manipulation of an independent variable occurs at different points along the horizontal axes of multiple-tiered graphs, a dog-leg connecting the condition change lines from one tier to the next makes it easy to follow the sequence and timing of events in the experiment (see Figure 6.20).

Unplanned events that occur during an experiment or treatment program, as well as minor changes in procedure that do not warrant a condition change line, can be indicated by placing small arrows, asterisks, or other symbols next to the relevant data points (see Figure 6.6) or just under the *x*-axis (see Figure 6.20). The figure caption should explain any special symbols.

Condition Labels

Labels identifying the conditions in effect during each period of an experiment are centered above the space delineated by the condition change lines. Whenever space permits, condition labels should be parallel to the horizontal axis. Labels should be brief but descriptive (e.g., *Contingent Praise* is preferable to *Treatment*), and the labels should use the same terms or phrases used in the accompanying text describing the condition. Abbreviations may be used when space or design limitations prohibit printing the complete label. A single condition label should be placed above and span across labels identifying minor changes within that condition (see Figure 6.19). Numbers are sometimes added to condition labels to indicate the number of times the condition has been in effect during the study (e.g., Baseline 1, Baseline 2).

Plot Data Points and Draw Data Paths

Data Points

When graphing data by hand, behavior analysts must take great care to ensure that they plot each data point exactly on the coordinate of the horizontal and vertical axis values of the measurement it represents. The inaccurate placement of data points is an unnecessary source of error in graphic displays, which can lead to mistakes in clinical judgment and/or experimental method. Accurate placement is aided by careful selection of graph paper with grid lines sized and spaced appropriately for the data to be plotted. When many different values must be plotted within a small distance on the vertical axis, a graph paper with many grid lines per inch should be used.

Should a data point fall beyond the range of values described by the vertical axis scale, it is plotted just above the scale it transcends, with the actual value of the measurement printed in parentheses next to the data point. Breaks in the data path leading to and from the off-the-scale data point also help to highlight its discrepancy (see Figure 6.20, Session 19).

Data points should be marked with bold symbols that are easily discriminated from the data path. When only one set of data is displayed on a graph, solid dots are most often used. When multiple data sets are plotted on the same set of axes, a different geometric symbol should be used for each set of data.

The symbols for each data set should be selected so that the value of each data point can be determined when data points fall near or on the same coordinates on the graph (see Figure 6.19, Sessions 9–11).

Data Paths

Data paths are created by drawing a straight line from the center of each data point in a given data set to the center of the next data point in the same set. All data points in a given data set are connected in this manner with the following exceptions:

- Data points falling on either side of a condition change line are not connected.
- Data points should not be connected across a significant span of time in which behavior was not measured. To do so implies that the resultant data path represents the level and trend of the behavior during the span of time in which no measurement was conducted.
- Data points should not be connected across discontinuities of time in the horizontal axis (see Figure 6.19, 1-week school vacation).
- Data points on either side of a regularly scheduled measurement period in which data were not collected or were lost, destroyed, or otherwise not available (e.g., participant's absence, recording equipment failure) should not be joined together (see Figure 6.20, Belva's graph, baseline condition).
- Follow-up or postcheck data points should not be connected with one another (see Figure 6.19) unless they represent successive measures spaced in time in the same manner as measures obtained during the rest of the experiment (see Figure 6.20).

When multiple data paths are displayed on the same graph, different styles of lines, in addition to different symbols for the data points, may be used to help distinguish one data path from another (see Figure 6.19). The behavior represented by each data path should be clearly identified, either by printed labels with arrows drawn to the data path (see Figures 6.19 and 6.20) or by a legend showing models of the symbols and line styles (see Figure 6.13). When two data sets travel the same path, their lines should be drawn close to and parallel with one another to help clarify the situation (see Figure 6.19, Sessions 9–11).

Write Figure Caption

Printed below the graph, the figure caption should give a concise but complete description of the figure. The caption should also direct the viewer's attention to any features of the graph that might be overlooked (e.g., scale changes) and should explain the meaning of any added symbols representing special events.

Print Graph

Graphs should be printed in only one color—black. Although the use of color can enhance the attractiveness of a visual display and highlight certain features, it is discouraged in the scientific presentation of data. Every effort must be made to let the data stand on their own. The use of color can encourage perceptions of performance or experimental effects that differ from perceptions of the same data displayed in black. The fact that graphs and charts may be reproduced in journals and books is another reason for using black only.

Constructing Graphs with Computer Software

Software programs for producing computer-generated graphs are available and becoming both increasingly sophisticated and easier to use. Most of the graphs displayed throughout this book were constructed with computer software. Even though computer graphics programs offer a tremendous time savings over hand-plotted graphs, careful examination should be made of the range of scales available and the printer's capability for both accurate data point placement and precise printing of data paths.

Most computer generated graphs are produced with Microsoft Excel software. Detailed task analyses and step-by-step instructions for creating behavioral graphs using Microsoft Excel are available (e.g., Cole & Witts, 2015; Deochand, 2017; Deochand, Costello, & Fuqua, 2015; Dixon et al., 2009; Lo & Konrad, 2007). Vanselow and Bourret (2012) developed an online interactive tutorial for creating graphs with Excel.

Software programs that automatically create graphs from digitally collected or practitioner-uploaded data save time that can be devoted to other clinical or instructional activities. Autographing software can help clinicians keep graphs of client performance current, enabling more frequent data-based decisions (e.g., Pinkelman & Horner, 2016; Sleeper et al., 2017).

INTERPRETING GRAPHICALLY DISPLAYED BEHAVIORAL DATA

The effects of an intervention that produces dramatic, replicable changes in behavior that last over time are readily seen in a well-designed graphic display. People with little or no formal training in behavior analysis can read the graph correctly in such cases. Many times, however, behavior changes are not so large, consistent, or durable. Behavior sometimes changes in sporadic, temporary, delayed, or seemingly uncontrolled ways; and sometimes behavior may hardly change at all. Graphs displaying these kinds of data patterns often reveal equally important and interesting subtleties about behavior and its controlling variables.

Behavior analysts employ a systematic form of examination known as **visual analysis** to interpret graphically displayed data. Visual inspection of graphs showing behavior change in relation to the presence or absence of environmental variables is the primary method of data analysis in behavior analysis. Visual analysis of data from an applied behavior analysis experiment is conducted to answer two questions: (a) Did behavior change in a meaningful way? and (b) if so, to what extent can that behavior change be attributed to the independent variable? Although there are no formalized rules for visual analysis, the dynamic nature of behavior, the scientific and technological necessity of discovering effective interventions, and the applied requirement of producing socially meaningful levels of performance all combine to focus the behavior analyst's interpretive attention on certain properties common to all behavioral data: (a) variability,

(b) level, and (c) trend. Visual analysis entails examining each of these characteristics both within and across the conditions and phases of an experiment.[12]

Before attempting to interpret the data displayed in a graph, the viewer should carefully examine the graph's overall construction. First, read the figure legend, axis labels, and all condition labels to obtain a basic understanding of what the graph is about. The viewer should then look at the scaling of each axis, taking note of the location, numerical value, and relative significance of any scale breaks. Did the graph maker truncate the vertical axis scale to emphasize behavior change?

> A common complaint of this is that it gives the appearance of severity when none exists. First, this is why charts have scales. Blaming a chart's creator for a reader who doesn't look at clearly labeled axes is like blaming a supermarket for selling someone food he's allergic to.
>
> Second, the degree to which the chart emphasizes certain aspects of the data is a judgment of storytelling not chart-making. Sure, this mindset can yield misleading displays, but how is that different than words? Charts should be fair, not impartial. (Yanofsky, 2015, n.p.)

Next, visually track each data path to determine that the data points are properly connected. Does each data point represent a single measurement or observation, or are the data "blocked" such that each data point represents an average or some other summary of multiple measurements? Do the data show the performance of an individual subject or the average performance of a group of subjects? If blocked or group data are displayed, is a visual representation of the range or variation of scores provided; or do the data themselves allow determination of the amount of variability that was collapsed in the graph? For example, if the horizontal axis is scaled in weeks and each data point represents a student's average score for a week of daily five-word spelling tests, data points falling near 0 or at the top end of the closed scale, such as 4.8, pose little problem because they can be the result of only minimal variability in the daily scores for that week. However, data points near the center of the scale, such as 2 to 3, can result from either stable or highly variable performance.

Only when satisfied that a graph is properly constructed and does not distort the events it represents, should the viewer examine the data to discover what they reveal about the behavior during each condition of the study. If the viewer suspects the graph's construction distorts, overstates, or hides important features, interpretive judgments of the data should be withheld (see Box 6.2).

BOX 6.2

Don't Be Fooled by a Pretty Graph

Graphic displays tell visual stories of the data they represent. But just as storytellers can conceal, distort, and mislead by selecting and sequencing biased words and phrases, graph makers can make certain interpretations of the data more or less likely by manipulating chart features. When looking at a graphic display, ask yourself whether it has been designed to tell a story that accurately reflects the underlying data, or whether it has been designed to tell a story more closely aligned with what the designer would like you to believe. Truncated or expanded axes with questionable scaling; extraneous color and depth cues; misleading labels, legends, and captions; and bar graphs with y-axes that do not include 0 are common forms of graphic foolery (Bergstrom & West, 2018a).

Why does a bar graph need to include 0 on the dependent axis, whereas a line graph need not do so? Our view is that the two types of graphs are telling different stories. By its design, a bar graph emphasizes the *absolute magnitude* of values associated with each category, whereas a line graph emphasizes the *change* in the dependent variable (usually the y value) as the independent variable (usually the x value) changes.

For a bar graph to provide a representative impression of the values being plotted, the visual weight of each bar—the amount of ink on the page, if you will—must be proportional to the value of that bar. Setting the axis above zero interferes with this. For example, if we create a bar graph with values 15 and 20 but set the axis at 10, the bar corresponding to 20 has twice the visual weight of the bar corresponding to 15, despite the value 20 being only 4/3 of the value 15. (Bergstrom & West, 2018b, n.p.)

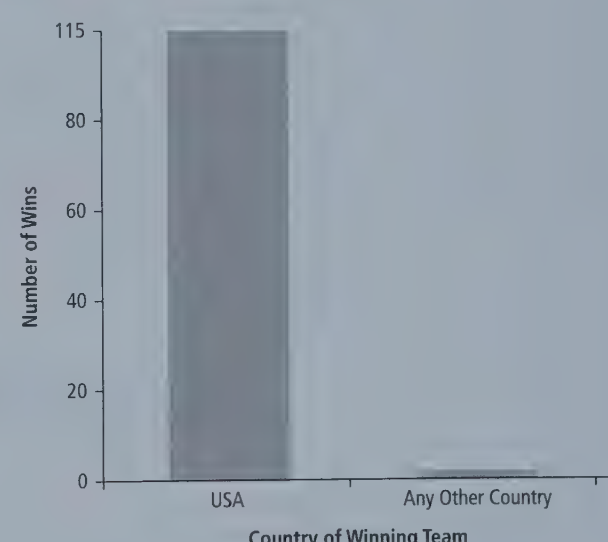


Figure A Country of Baseball's World Series–Winning Teams (1903–2017).

Don't be fooled by graphs comparing data sets that have no logical or possible relation to one another. Figure A, for example, compares the number of baseball teams located

(continued)

BOX 6.2 (continued)

in the United States who have won baseball's World Series with the number of World Series winners from outside the United States. The story told by this graph is U.S. dominance over the rest of the baseball-playing world. Although the data depicted are accurate, they represent dramatically different opportunities. From the first World Series in 1903 through 1968 there were 0 non-U.S. major league teams. The Montreal Expos competed in the National League from 1969 to 2004. The Toronto Blue Jays, World Series winners in 1992 and 1993, have played in the American League since 1977.

Comparing multiple data sets with different scales for each set can create even more deceptive graphs. One must be especially careful comparing multiple data series on the graph, with different scales for each series. Figure B shows a correlation between the use of the herbicide glyphosate (trade name Roundup) on corn and soybean crops and the increasing incidence of cancer. The authors reported: "We found strong correlations for cancers of the liver, kidney, bladder/urinary and thyroid. . . . Thyroid and bladder cancers especially seem to track with the advent of GE [genetically engineered] crops and associated glyphosate applications" (Swanson, Leu, Abrahamson, & Wallet, 2014, p. 16, words in brackets added).

Exposure to Roundup may well have serious health consequences, but whatever they may be, this particular graph is not persuasive. First of all, there's the obvious point that correlation is not causation. One would find a similar correlation between cell phone usage and hypertension, for example—or even between cell phone usage and Roundup usage! The authors make no causal claims, but we fail to see the value in fishing for correlations.

The left-hand axis doesn't go to zero. We've already noted why this is problematic. But it gets worse. Both the scale and *the intercept* of the other vertical axis, at right, have been adjusted so that the red curve traces the peaks of the yellow bars. Most remarkably, to make the curves do this, the axis has to go all the way to *negative* 10% GE corn planted and *negative* 10,000 tons glyphosate used! We've noted that the *y*-axis need not go to zero, but if it goes to a negative value for a quantity (percentage or tonnage) that can only take on positive values, this should set off alarm bells. (Bergstrom & West, 2018b, emphasis in original, n.p.)

While a challenge to everyone, detecting graphic sleight of hand is particularly important for applied behavior analysts, who rely on visual analysis of graphed data to determine the story of behavior change. Fortunately, transparency is a characteristic of behavior analytic research and practice. Many graphic displays used in behavior analysis provide direct access to the original data, which allows the inquisitive or skeptical viewer to replot the data on a new set of axes.

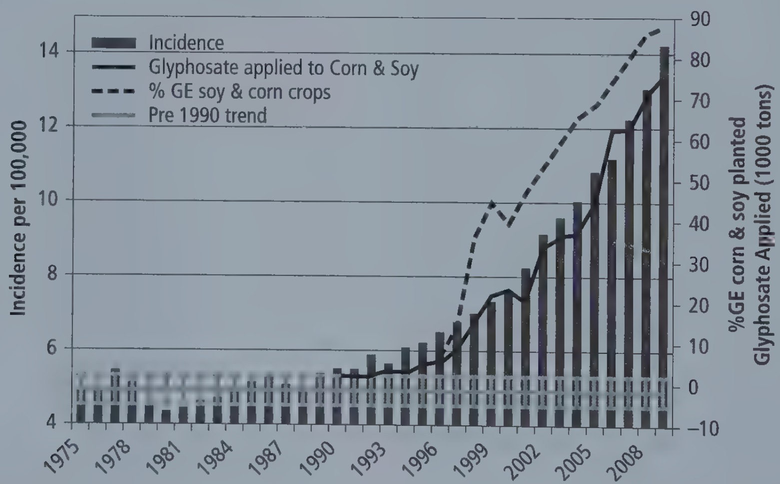


Figure B Graph suggesting correlation between rising incidence of thyroid cancer and use of herbicide glyphosate on U.S. corn and soy crops that are genetically engineered.

From "Genetically Engineered Crops, Glyphosate and the Deterioration of Health in the United States of America," by N. L. Swanson, A. Leu, J. Abrahamson, & B. Wallet, 2014, *Journal of Organic Systems, 9*, p. 18. Copyright 2014 by *Journal of Organic Systems*. Used by permission.

Visual Analysis Within Conditions

Data within a given condition are examined to determine (a) the number of data points, (b) the nature and extent of variability in the data, (c) the absolute and relative level of the behavioral measure, and (d) the direction and degree of any trend(s) in the data.[13]

Number of Data Points

First, the viewer should determine the quantity of data reported during each condition. This entails a simple counting of data points. As a general rule, the more measurements of the dependent variable per unit of time and the longer the period of time in which measurement occurred, the more confidence one can have in the data

path's estimation of the true course of behavior change (given, of course, a valid and accurate observation and measurement system).

The number of data points needed to provide a believable record of behavior during a given condition also depends on how many times the same condition has been repeated during the study. As a rule, fewer data points are needed in subsequent replications of an experimental condition if the data depict the same level and trend in performance that were noted in earlier applications of the condition.

The published literature of applied behavior analysis also plays a part in determining how many data points are sufficient. In general, briefer phases are required of experiments investigating relations between previously studied and well-established variables if the results are also similar to those of the previous studies. More data are needed to demonstrate new findings, whether or not new variables are under investigation.

There are other exceptions to the rule of the-more-data-the-better. Ethical concerns do not permit the repeated measurement of certain behaviors (e.g., self-injurious behavior) under an experimental condition in which there is little or no expectation for improvement (e.g., during a no-treatment baseline condition or a condition intended to reveal variables that exacerbate problem behavior). Also, there is little purpose in repeated measurement in situations in which the subject cannot logically perform the behavior (e.g., measuring the number of correct answers to long division problems when concurrent observations indicate that the student has not learned the necessary component skills of multiplication and subtraction). Nor are many data points required to demonstrate that behavior did not occur when in fact it had no opportunity to occur.

Familiarity with the response class measured and the conditions under which it was measured may be the graph viewer's biggest aid in determining how many data points constitute believability. The quantity of data needed in a given condition or phase is also partly determined by the analytic tactics employed in a given study. Experimental design tactics are described in Chapters 7 through 10.

Variability

How often and the extent to which repeated measures of behavior yield different outcomes is called **variability**. A high degree of variability (or "bounce") within a given condition usually indicates that the researcher or practitioner has achieved insufficient control over the factors influencing the behavior. (An important exception is when the purpose of an intervention is to produce a high degree of variability.) In general, the greater the variability within a given condition, the greater the number of data points that are necessary to establish a predictable pattern of performance. By contrast, fewer data points are required to present a predictable pattern of performance when those data reveal relatively little variability.

Practitioners and researchers should also look for variability in the form of cyclical patterns, such as those shown in the bottom panel of Figure 6.1.

Level

The value on the vertical axis scale around which a set of behavioral measures converge is called **level**. In the visual analysis of behavioral data, level is examined within a condition in terms of its absolute value on the *y*-axis (mean, median, and/or range), the degree of stability or variability, and the extent of change from one level to another. The graphs in Figure 6.21 illustrate four different combinations of level and variability.

The mean level of a series of behavioral measures within a condition can be graphically illustrated by the addition of a *mean level line*: a dashed horizontal line drawn through a series of data points within a condition at that point on the vertical axis equaling the average value of the series of measures. Although mean level lines provide an easy-to-see summary of average performance within a given condition or phase, they should be used and interpreted with caution. With highly stable data paths, mean level lines pose no serious drawbacks. However, the less variability there is within a series of data points, the less need there is for a mean level line. For instance, a mean level line would serve little

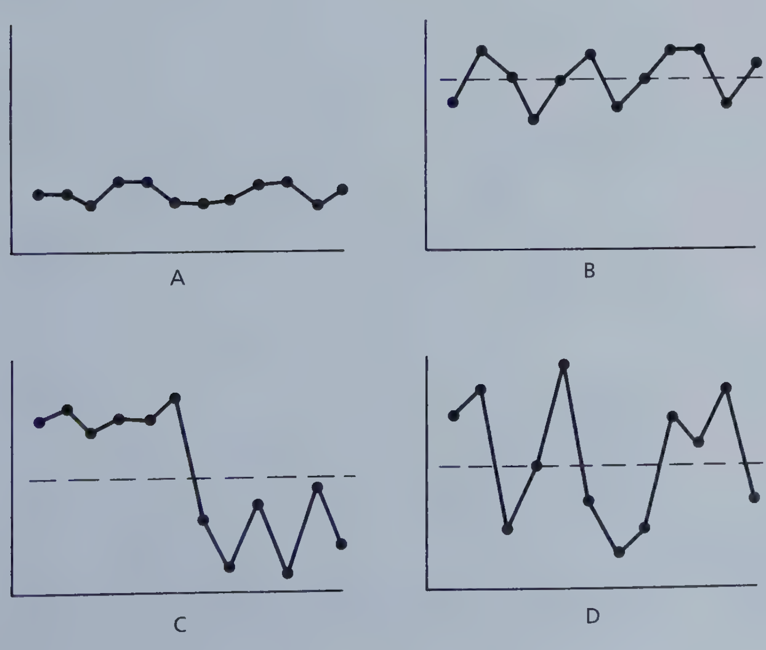


Figure 6.21 Four data paths illustrating (A) a low, stable level of responding; (B) a high, variable level of responding; (C) an initially high, stable level of responding followed by a lower, more variable level of responding; and (D) an extremely variable pattern of responding not indicative of any overall level of responding. Dashed horizontal lines on graphs B, C, and D represent the mean levels of responding.

purpose in Graph A in Figure 6.21. And although mean level lines have been added to Graphs B, C, and D in Figure 6.21, Graph B is the only one of the three for which a mean level line provides an appropriate visual summary of level. The mean level line in Graph C is not representative of any measure of behavior taken during the phase. The data points in Graph C show a behavior best characterized as occurring at two distinct levels during the condition and beg for an investigation of the factor(s) responsible for the clear change in levels. The mean level line in Graph D is inappropriate because the variability in the data is so great that only 4 of the 12 data points fall close to the mean level line.

A *median level line* is another method for visually summarizing the overall level of behavior in a condition. Because a median level line represents the most typical performance within a condition, it is not so influenced by one or two measures that fall far outside the range of the remaining measures. Therefore, one should use a median level line instead of a mean level line to graphically represent the central tendency of a series of data points that include several outliers, either high or low.

Change in level within a condition is determined by calculating the difference in absolute values on the *y*-axis between the first and last data points within the condition. Another method, somewhat less influenced by variability in the data, is to compare the difference between the median value of the first three data points in the condition with the median value of the final three data points in the condition (Koenig & Kunzelmann, 1980).

Trend

The overall direction taken by a data path is its **trend**. Trends are described in terms of their direction (increasing, decreasing, or zero trend), degree or magnitude, and extent of variability of data points around the trend. The graphs in Figure 6.22 illustrate a variety of trends. The direction and degree of trend in a series of graphed data points can be visually represented by a straight line drawn through the data. called a *trend line* or *line of progress*. Several methods for calculating and fitting trend lines to a series of data have been developed. One can simply inspect the graphed data and draw a straight line that visually provides

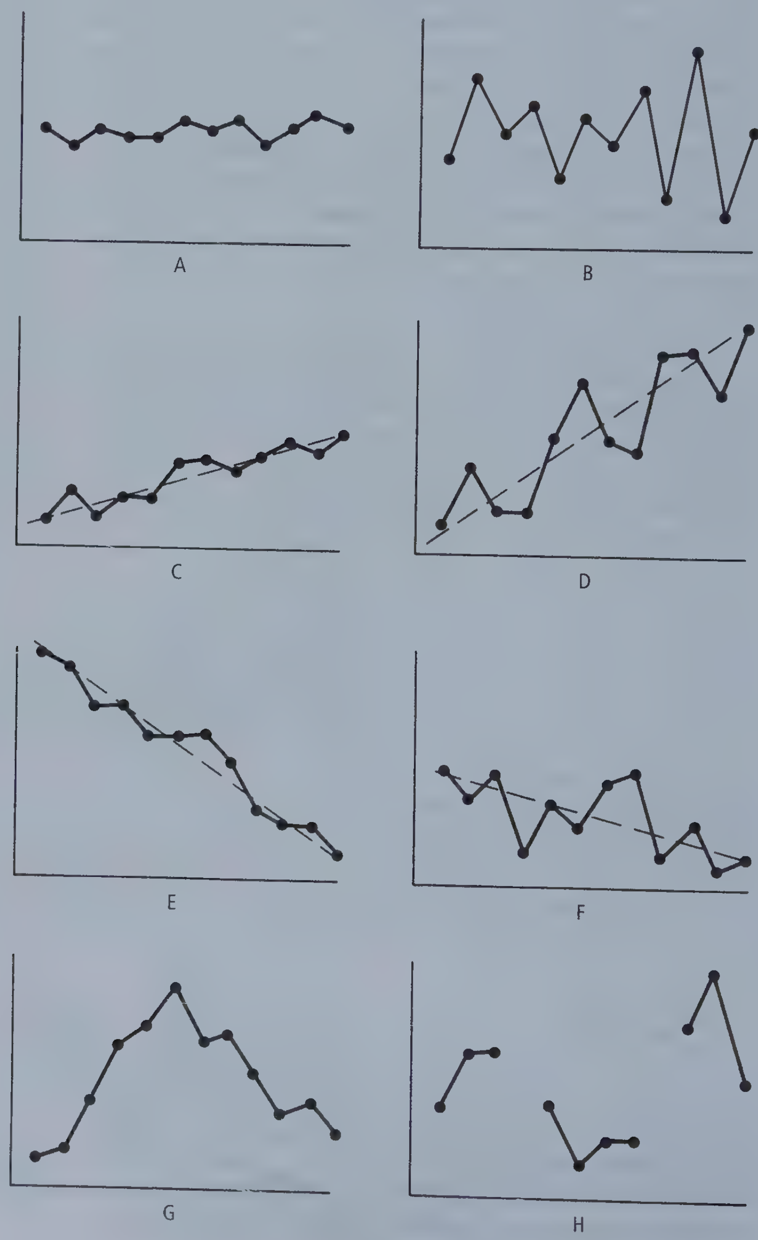


Figure 6.22 Data patterns indicating various combinations of trend direction, degree, and variability: (A) zero trend, high stability; (B) zero trend, high variability; (C) gradually increasing stable trend; (D) rapidly increasing variable trend; (E) rapidly decreasing stable trend; (F) gradually decreasing variable trend; (G) rapidly increasing trend followed by rapidly decreasing trend; (H) no meaningful trend, too much variability and missing data. Split-middle lines of progress have been added to Graphs C–F.

the best fit through the data. For this freehand method, Lindsley (1985) suggested ignoring one or two data points that fall well beyond the range of the remaining values in a data series and fitting the trend line to the remaining scores. Although the freehand method is the fastest way of drawing trend lines and can be useful for the viewer of a published graph, hand-drawn trend lines may not always result in an accurate representation of trend and are typically not found in graphs of published studies.

Trend lines can also be calculated using a mathematical formula called the ordinary least-squares linear regression equation (McCain & McCleary, 1979; Parsonson & Baer, 1978). Trend lines determined in this fashion have the advantage of complete reliability: The same trend line will always result from the same data set. The disadvantage of this method is the many mathematical operations that must be performed to calculate the trend line. A computer program that can perform the equation eliminates the time concern in calculating a least-squares trend line.

A method of calculating and drawing lines of progress that is more reliable than the freehand method and much less time consuming than linear regression methods is the **split-middle line of progress**. White (2005) developed the split-middle technique for use with rate data plotted on semilogarithmic charts, and it has proven a useful technique for predicting future behavior from such data. Split-middle lines of progress can also be drawn for data plotted against an equal-interval vertical axis, but it must be remembered that such a line is only an estimate that summarizes the overall trend (Bailey, 1984). A trend line cannot be drawn by any method through a series of data points spanning a scale break in the vertical axis and generally should not be drawn across scale breaks in the horizontal axis.

The specific degree of acceleration or deceleration of trends in data plotted on semilogarithmic charts can be quantified in numerical terms. For example, on the daily Standard Celeration Chart, a "times-2" celeration means that the response rate is doubling each week, and a "times-1.25" means that the response rate is accelerating by a factor of one fourth each week. A "divide-by-2" celeration means that each week the response rate will be one half of what it was the week before, and a "divide-by-1.5" means that the frequency is decelerating by one third each week.

There is no direct way to determine visually from data plotted on equal-interval charts the specific rates at which trends increase or decrease. But visual comparison of trend lines drawn through data on equal-interval charts can provide important information about the relative rates of behavior change.

A trend may be highly stable with all of the data points falling on or near the trend line (see Figure 6.22, Graphs C and E). Data paths can also follow a trend even though a high degree of variability exists among the data points (see Figure 6.22, Graphs D and F).

Visual Analysis Between Conditions

After inspection of the data within each condition or phase of a study, visual analysis proceeds with a comparison of data between conditions. Drawing proper conclusions entails comparing the previously discussed properties of behavioral data—level, trend, and variability between different conditions and among similar conditions.

A condition change line indicates that an independent variable was manipulated at a given point in time. To determine whether an immediate change in behavior occurred at that point in time, one needs to examine the difference between the last data point before the condition change line and the first data point in the new condition.

The data are also examined in terms of the overall level of performance between conditions. In general, when all data points in one condition fall outside the range of values for all data points in an adjacent condition (i.e., there is no overlap of data points between the highest values obtained in one condition and the lowest values obtained in the other condition), there is little doubt that behavior changed from one condition to the next. When many data points in adjacent conditions overlap one another on the vertical axis, less confidence can be placed in the effect of the independent variable associated with the change in conditions.[14]

Mean or median level lines can be helpful in examining the overall level between conditions. However, using mean or median level lines to summarize and compare the overall central tendency of data across conditions poses two serious problems. First, the viewer of such a visual display must guard against letting "apparently large differences among measures of central tendency visually overwhelm the presence of equally large amounts of uncontrolled variability" (Johnston & Pennypacker, 1980, p. 351). Emphasizing mean changes in performance in a graphic display can suggest a greater degree of experimental control was obtained than is warranted by the data. In the top graph of Figure 6.23, half of

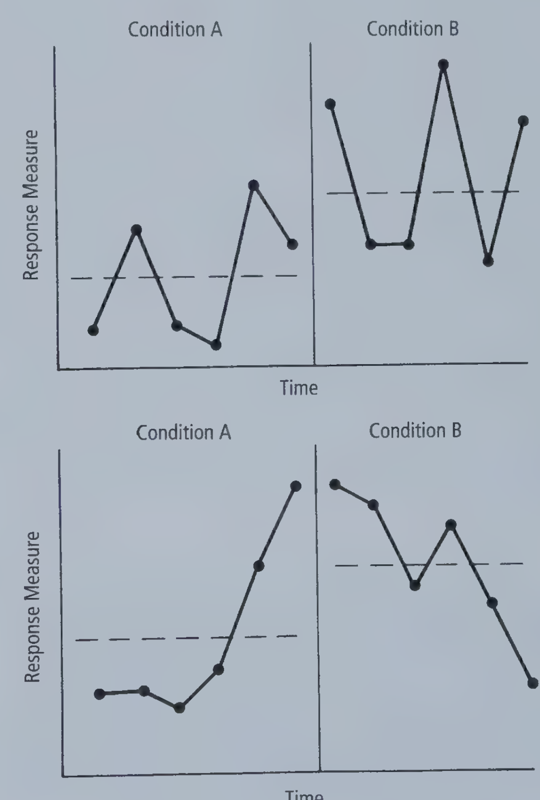


Figure 6.23 Inappropriate use of mean level lines, encouraging interpretation of a higher overall level of responding in Condition B when extreme variability (top graph) and trends (bottom graph) warrant different conclusions.

the data points in Condition B fall within the range of values of the measures taken during Condition A, but the mean level lines suggest a clear change in behavior. Second, measures of central tendency can obscure important trends in the data that warrant interpretations other than those suggested by the central tendency indicators. Although a mean or median line accurately represents the average or typical performance, neither provides any indication of increasing or decreasing performance. In the bottom graph of Figure 6.23, for example, the mean line suggests a higher level of performance in Condition B than in Condition A, but an examination of trend yields a very different picture of behavior change within and between Conditions A and B.

Those analyzing behavioral data should also note any changes in level that occur after a new condition has been in place for some time and any changes in level that occur early in a new condition but are later lost. Such delayed or temporary effects can indicate that the independent variable must be in place for some time before behavior changes, or that the temporary level change was the result of an uncontrolled variable. Either case calls for further investigation in an effort to isolate and control relevant variables.

Visual analysis of data between adjacent conditions includes an examination of the trends exhibited by the data in each condition to determine whether the trend found in the first condition changed in direction or slope during the subsequent condition. In practice, because each data point in a series contributes to level and trend, the two characteristics are viewed in conjunction with one another. Figure 6.24 presents stylized data paths illustrating four basic combinations of change or lack of change in level and trend between adjacent conditions. Of course, many other data patterns could display the same characteristics. Idealized, straight-line data paths that eliminate the variability found in most repeated measures of behavior have been used to highlight level and trend.

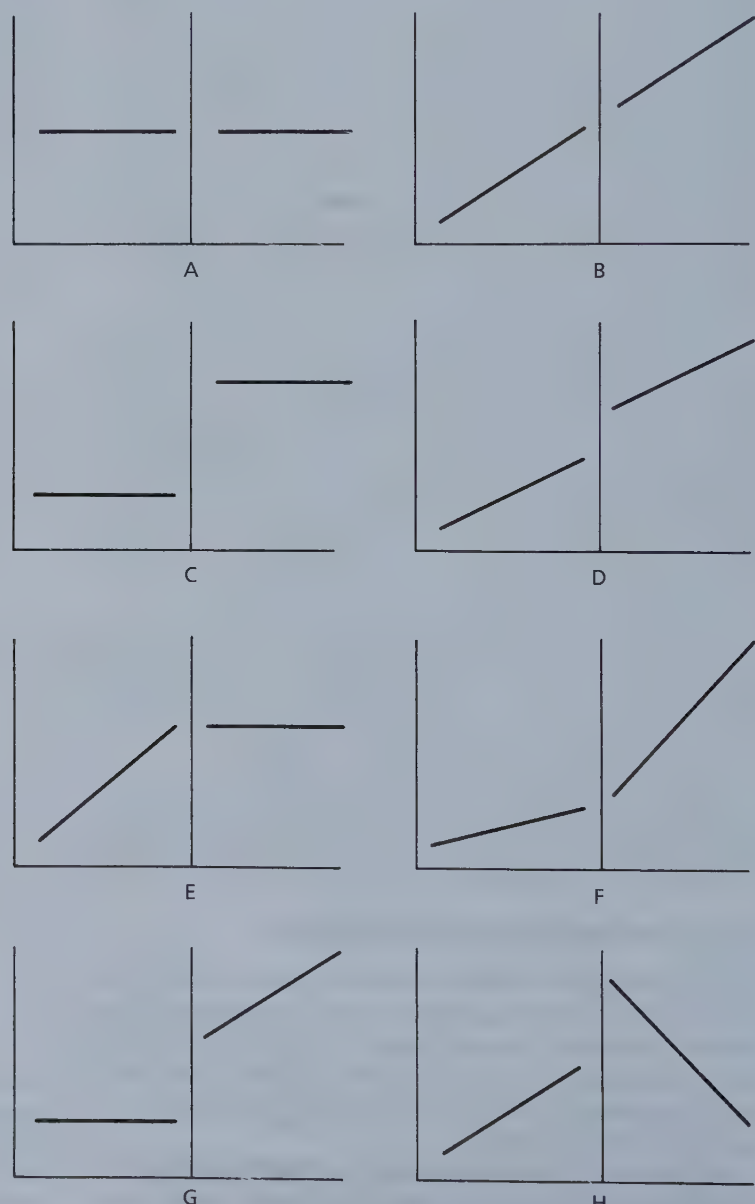


Figure 6.24 Stylized data paths illustrating the different combinations of change or lack of change in level and trend between two adjacent conditions: Graphs A and B show no change in either level or trend between the two conditions, Graphs C and D show changes in level and no change in trend, Graphs E and F depict no immediate change in level and a change in trend, and Graphs G and H reveal change in both level and trend.

From "Time-Series Analysis in Operant Research," by R. R. Jones, R. S. Vaught, and M. R. Weinrott, 1977, *Journal of Applied Behavior Analysis*, 10, p. 157. Copyright 1977 by the Society for the Experimental Analysis of Behavior, Inc. Adapted by permission.

Visual analysis includes not only an examination and comparison of changes in level and trend between adjacent conditions, but also an examination of performance across similar conditions. Interpreting what the data from an applied behavior *analysis* mean requires more than visual analysis and the identification and description of level, trend, and variability. When behavior change is demonstrated over the course of a treatment program or research study, the next question to be asked is, Was the change in behavior a function of the treatment or experimental variable? The remaining chapters of Part Three describe strategies and tactics of experimental design used by applied behavior analysts in an effort to provide a meaningful answer.

SUMMARY

1. Applied behavior analysts document and quantify behavior change by direct and repeated measurement of behavior, and the product of those measurements is called *data*.
2. Graphs are relatively simple formats for visually displaying relationships among and between a series of measurements and relevant variables.

Purpose and Benefits of Graphic Displays of Behavioral Data

3. Graphing each measure of behavior as it is collected provides the practitioner or researcher with an immediate and ongoing visual record of the participant's behavior, allowing treatment and experimental decisions to be responsive to the participant's performance.
4. Direct and continual contact with the data in a readily analyzable format enables the practitioner or researcher to identify and investigate interesting variations in behavior as they occur.
5. As a judgmental aid for interpreting experimental results, graphic display is a fast, relatively easy-to-learn method that imposes no arbitrary levels of significance for evaluating behavior change.
6. Visual analysis of graphed data is a conservative method for determining the significance of behavior change; only variables able to produce meaningful effects repeatedly are considered significant, and weak and unstable variables are screened out.
7. Graphs enable and encourage independent judgments of the meaning and significance of behavior change by others.
8. Graphs can serve as effective sources of feedback to the people whose behavior they represent.

Graphs Used by Applied Behavior Analysts

9. Line graphs, the most commonly used format for the graphic display of behavioral data, are based on the Cartesian plane, a two-dimensional area formed by the intersection of two perpendicular lines.
10. Major parts of the simple line graph are the horizontal axis (also called the x-axis), the vertical axis (also called the y-axis), condition change lines, condition labels, data points, the data path, and the figure caption.
11. Graphs with multiple data paths on the same set of axes are used in applied behavior analysis to show (a) two or more dimensions of the same behavior, (b) two or more different behaviors, (c) the same behavior under different and alternating experimental conditions, (d) changes in target behavior relative to the changing values of an independent variable, and (e) the behavior of two or more participants.
12. A second vertical axis, which is drawn on the right-hand side of the horizontal axis, is sometimes used to show different scales for multiple data paths.
13. Bar graphs are used for two primary purposes: (a) to display discrete data not related by an underlying dimension that can be used to scale the horizontal axis and (b) to summarize and enable easy comparison of the performance of a participant or group of participants during the different conditions of an experiment.
14. Each data point on a cumulative record represents the total number of responses emitted by the subject since measurement began. The steeper the slope of the data path on a cumulative graph, the higher the response rate.
15. Overall response rate refers to the average rate of response over a given time period; a local response rate refers to the rate of response during a smaller period within a larger period for which an overall response rate has been given.
16. Cumulative records are especially effective for displaying data when (a) the total number of responses made over time is important, (b) the graph is used as a source of feedback to the subject, (c) the target behavior can occur only once per measurement period, and (d) a fine analysis of a single instance or portions of data from an experiment is desired.
17. Ratio charts use a logarithmic-scaled y-axis so that changes in behavior that are of equal proportion (e.g., doublings of the response measure) are represented by equal distances on the vertical axis.
18. The Standard Celeration Chart is a six-cycle multiply-divide graph that enables the standardized charting of celeration, a linear measure of frequency change across time, a factor by which frequency multiplies or divides per unit of time.
19. A scatterplot shows the relative distribution of individual measures in a data set with respect to the variables depicted by the x- and y-axes.

Constructing Line Graphs

20. The vertical axis is drawn to a length approximately two-thirds that of the horizontal axis.

21. The horizontal axis is marked in equal intervals, each representing, from left to right, the chronological succession of equal time periods within which behavior was measured.

22. Scale breaks on the horizontal axis indicate discontinuities of time.

23. The vertical axis is scaled relative to the dimension of behavior measured, the range of values of the measures obtained, and the social significance of various levels of change in the target behavior.

24. Condition change lines indicate changes in the treatment program or manipulations of an independent variable and are drawn to the same height as the vertical axis.

25. A brief, descriptive label identifies each condition of an experiment or behavior change program.

26. Data points should be accurately placed with bold, solid dots. When multiple data paths are used, different geometric symbols are used to distinguish each data set.

27. Data paths are created by connecting successive data points with a straight line.

28. Successive data points should not be connected when (a) they fall on either side of a condition change line; (b) they span a significant period in which behavior was not measured; (c) they span discontinuities of time on the horizontal axis; (d) they fall on either side of a regularly scheduled measurement period in which data were not collected or were lost, destroyed, or otherwise not available; (e) they fall in a follow-up or postcheck period that is not regularly spaced in time in the same manner as the rest of the study; or (f) one member of the pair falls outside the range of values described by the vertical axis.

29. The figure caption provides a concise but complete description of the graph, giving all of the information needed to interpret the display.

30. Graphs should be printed in black ink only.

Interpreting Graphically Displayed Behavioral Data

31. Visual analysis of graphed data attempts to answer two questions: (a) did a socially meaningful change in behavior take place and (b) if so, can the behavior change be attributed to the independent variable?

32. Before beginning to evaluate the data displayed in a graph, a careful examination of the graph's construction should be undertaken. If distortion is suspected from the features of the graph's construction, the data should be replotted on a new set of axes before interpretation is attempted.

33. Blocked data and data representing the average performance of a group of subjects should be viewed with the understanding that significant variability may have been lost in the display.

34. Visual analysis of data within a given condition focuses on the number of data points, the variability of performance, the level of performance, and the direction and degree of any trends in the data.

35. As a general rule, the more data in a condition and the greater the stability of those data, the more confidence one can place in the data path's estimate of behavior during that time. The more variability in the behavioral measures during a condition, the greater the need for additional data.

36. *Variability* refers to the frequency and degree to which multiple measures of behavior yield different outcomes. A high degree of variability within a given condition usually indicates that little or no control has been achieved over the factors influencing the behavior.

37. *Level* refers to the value on the vertical axis around which a series of data points converges. When the data in a given condition all fall at or near a specific level, the behavior is considered stable with respect to level; to the extent that the behavioral measures vary considerably from one to another, the data are described as showing variability with respect to level. In cases of extreme variability, no particular level of performance is evidenced.

38. Mean or median level lines are sometimes added to graphic displays to represent the overall average or typical performance during a condition. Mean and median level lines should be used and interpreted with care because they can obscure important variability and trends in the data.

39. *Trend* refers to the overall direction taken by a data path; trends are described in terms of their direction (increasing, decreasing, or zero trend), degree (gradual or steep), and the extent of variability of data points around the trend.

40. Trend direction and degree can be visually represented by drawing a trend line, or line of progress, through a series of data points. Trend lines can be drawn freehand, using the least-squares regression equation, or using a method called the split-middle line of progress. Split-middle lines of progress can be drawn quickly and reliably and have proven useful in analyzing behavior change.

41. Visual analysis of data across conditions determines whether change in level, trend, and/or variability occurred and to what extent any changes were significant.

KEY TERMS

bar graph
cumulative record
cumulative recorder
data
data path
dependent variable
equal-interval scale
graph
independent variable
level
line graph
local response rate
overall response rate
precision teaching
ratio scale
scatterplot
split-middle line of progress
Standard Celeration Chart
trend
variability
visual analysis

MULTIPLE-CHOICE QUESTIONS

1. Visual analysis is considered a(n) ________ method for determining the social significance of behavior change.
   a. Temporary
   b. Arrangement
   c. Subjective
   d. Conservative
   Hint: (See "Purpose and Benefits of Graphic Displays of Behavioral Data")
2. Graphic displays of behavioral data are considered simple formats for visually displaying ________ among and between a series of measurements and relevant variables.
   a. Time series
   b. Data
   c. Relationships
   d. Graphs
   Hint: (See "Purpose and Benefits of Graphic Displays of Behavioral Data")
3. This type of graph is also known as a histogram and is useful in summarizing behavioral data allowing for a quick comparison of performance across participants and/or conditions.
   a. Line graph
   b. Scatterplot
   c. Cumulative Record
   d. Bar graph
   Hint: (See "Graphs Used by Applied Behavior Analysts")
4. This is the most commonly used graph in applied behavior analysis.
   a. Line graph
   b. Scatterplot
   c. Cumulative Record
   d. Bar graph
   Hint: (See "Graphs Used by Applied Behavior Analysts")
5. You are interested in the total number of sight words accurately read since your data collection began last week. Which graphing convention allows you to most efficiently answer your inquiry?
   a. Line graph
   b. Scatterplot
   c. Cumulative Record
   d. Bar graph
   Hint: (See "Graphs Used by Applied Behavior Analysts")
6. You are interested in displaying the mean score for both your baseline and experimental condition. For this summary, you are willing to sacrifice the presentation and detection of variability and trends.
   a. Line graph
   b. Scatterplot
   c. Cumulative Record
   d. Bar graph
   Hint: (See "Graphs Used by Applied Behavior Analysts")
7. A ________ is depicted on a line graph by an open spot in the axis with a squiggly line at each end.
   a. Condition line
   b. Data set
   c. Scale break
   d. Horizontal axis
   Hint: (See "Line Graphs")
8. This part of a line graph is printed below the graph and gives a concise but complete description of the figure.
   a. Figure caption
   b. Data set
   c. Scale break
   d. Condition label
   Hint: (See "Line Graphs")

9. A ________ chart is a type of semilogarithmic chart useful for charting accelerating and decelerating performances over time.

   a. Cumulative time
   b. Standard celebration
   c. Scatter plot
   d. Bar graph

   Hint: (See "Graphs Used by Applied Behavior Analysts")

10. In order to produce a visual representation of an overall rate on a cumulative graph, the first and last data points of a given series of observations should be:

   a. Plotted bilaterally
   b. Plotted using different symbols
   c. Drawn simultaneously
   d. Connected with a straight line

   Hint: (See "Graphs Used by Applied Behavior Analysts")

11. The ________ of each data path on a cumulative record represents the different rates of acquisition.

   a. Position
   b. Slope
   c. Scatter
   d. Slant

   Hint: (See "Graphs Used by Applied Behavior Analysts")

12. A scatterplot shows the ________ distribution of individual measures in a data set with respect to the variables depicted by the both the x- and y-axes.

   a. Exact
   b. Scattered
   c. Relative
   d. Standard

   Hint: (See "Graphs Used by Applied Behavior Analysts")

13. Scatterplots are useful when applied behavior analysts are interested in finding out the ________ distribution of the target behavior. In other words, a scatterplot illustrates whether the behavior's occurrence is typically associated with certain time periods.

   a. Temporal
   b. Independent
   c. Dependent
   d. False

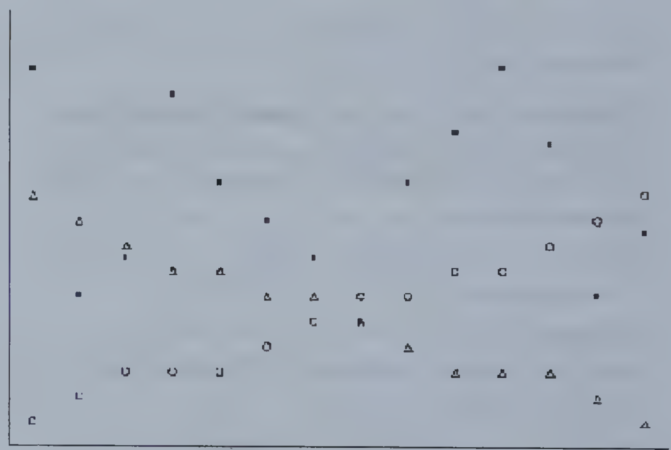
   Hint: (See "Graphs Used by Applied Behavior Analysts")

14. Which part of the line graph is the arrow pointed at?

   a. X-axis
   b. Condition change line
   c. Y-axis
   d. Data path

   Hint: (See "Line Graphs")

15. The type of graph illustrated in this example has disconnected data points and is used primarily to depict temporal distributions and reveal correlations between variables depicted on the y-axis with variables depicted on the x-axis.



   a. Cumulative record
   b. Line graph
   c. Bar graph
   d. Scatterplot

   Hint: (See "Line Graphs")

16. You are interested in beginning a self-management program to help you save money for a new car. Which type of graph would provide the most appropriate feedback and display your progress toward your goal?

   a. Cumulative record
   b. Line graph
   c. Bar graph
   d. Scatterplot

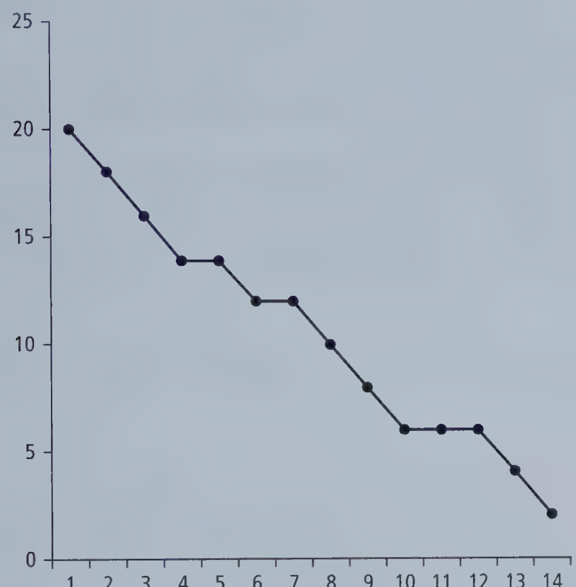
   Hint: (See "Graphs Used by Applied Behavior Analysts")

17. You are interested in comparing the effects of two different interventions on a single target behavior. You have decided to present each of the interventions in an alternating fashion and measure the changes in the value of your dependent variable. You would like to plot the results of both interventions on one graph. Which would be the best format to display your results?

   a. Cumulative record
   b. Line graph
   c. Bar graph
   d. Scatterplot

   Hint: (See "Graphs Used by Applied Behavior Analysts")

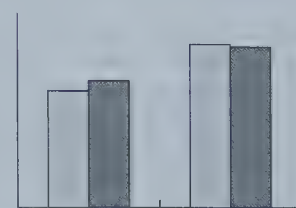
18. Which of the following best describes the direction, degree, and variability of the data path depicted in this graph?



   a. Rapidly increasing, variable trend
   b. Zero trend
   c. Gradually decreasing, variable trend
   d. Rapidly decreasing, stable trend

   Hint: (See "Interpreting Graphically Displayed Behavioral Data")

19. Which of the following best describes the direction, degree, and variability of the data path depicted in this graph?



   a. Rapidly increasing, variable trend
   b. Cannot be determined
   c. Gradually decreasing, variable trend
   d. Rapidly decreasing, stable trend

   Hint: (See "Interpreting Graphically Displayed Behavioral Data")

ESSAY-TYPE QUESTIONS

1. State a purpose and a benefit of using graphic displays of behavioral data for decision-making over listing a series of numbers in a tabular format.

   Hint: (See "Purpose and Benefits of Graphic Displays of Behavioral Data")

2. Describe a benefit of using visual analysis over statistical analysis in interpreting the social significance of behavior change.

   Hint: (See "Purpose and Benefits of Graphic Displays of Behavioral Data")

3. List and describe the parts of a basic line graph.

   Hint: (See "Parts of a Basic Line Graph")

4. Discuss why a line graph is often the most appropriate format for communicating changes in level, trend, and the variability of behavioral data when comparing the effects of intervention A to intervention B.

   Hint: (See "Graphs Used by Applied Behavior Analysts")

5. Discuss a situation in which a cumulative record would be preferable to a basic line graph.

   Hint: (See "Graphs Used by Applied Behavior Analysts")

6. Discuss a situation in which a scatterplot would be the most appropriate graphic format.

   Hint: (See "Graphs Used by Applied Behavior Analysts")

7. Which type of graph would you select to communicate changes in behavior when the behavior can only occur or not occur one time during a measurement period?

   Hint: (See "Graphs Used by Applied Behavior Analysts")

8. Define the following fundamental properties of behavior change: variability, level, and trend.

   Hint: (See "Interpreting Graphically Displayed Behavioral Data")

NOTES

1. Although often used as a singular construction (e.g., "The data *shows* that . . ."), *data* is the plural form of *datum*, noun of Latin origin; thus, *data* is used correctly with plural verbs (e.g., "These data *are* . . .").
2. Skillful visual analysis is not quite so simple, as will become evident later in this chapter and in Chapter 10.
3. A comparison of the visual analysis of graphed data and inferences based on statistical tests of significance is presented in Chapter 10.
4. Graphs, like statistics, can also be manipulated to make certain interpretations of the data more or less likely. Unlike statistics, however, most graphic displays used in behavior analysis provide direct access to the original data, which allows the inquisitive or doubtful reader to regraph the data and have another look.
5. The terms *condition* and *phase* are related but not synonymous. Properly used, *condition* indicates the environmental arrangements in effect; *phase* refers to a period within a study or behavior change program.
6. Functional communication training is described in Chapter 26.
7. To learn the interesting history and role of the cumulative recorder in the development of behavior analysis, see Lattal (2004) and Morris and Smith (2004).
8. Technically, Figure 6.12 does not represent true rates of response because the number of words spelled correctly was measured and not the rate, or speed, at which they were spelled. However, the slope of each data path represents the "rate" of mastering the spelling words in each session within the context of a total of 10 new words presented per session.
9. For detailed explanations of the Standard Celeration Chart and its use by precision teachers see Cooper, Kubina, and Malanga (1998); Graf and Lindsley (2002); Johnson and Street (2014); Kubina and Yurich 2012); Lindsley (1990, 1992, 1996); and Pennypacker, Gutierrez, and Lindsley (2003).
10. Perhaps the ultimate example of combining visual display techniques is Charles Minard's use of space-time-story graphics to illustrate the interrelations of six variables during Napoleon's ill-fated Russian campaign of 1812–1813 (see Tufte, 1983, p. 41). Tufte called Minard's graph perhaps "the best statistical graphic ever drawn"(p. 40).
11. Although most graphs published in behavior analysis journals since the mid-1990s and many graphs used by practitioners are constructed with computer software, we believe it is essential that applied behavior analysts know how to construct graphs by hand. The knowledge and skills required to create proper graphs by hand enable effective use of graphing software (cited later in this chapter).
12. Several studies have evaluated methods for training visual data analysis skills to behavior analysis students and practitioners (e.g., Kahng et al., 2010; Stewart, Carr, Brandt, & McHenry, 2007; Vanselow, Thompson, & Karsina, 2011; Wolfe & Slocum, 2015; Young & Daly, 2016).
13. When each data point represents multiple measures, a within-session analysis of the data may be warranted to detect behavioral functions and trends not apparent in aggregated data displays (Fahmie & Hanley, 2008). See Figure 6.14 for an example.
14. Whether a documented change in behavior should be interpreted as a function of the independent variable depends on the experimental design used in the study. Strategies and tactics for designing experiments are presented in Chapters 7 through 10.

CHAPTER 7

Analyzing Behavior Change: Basic Assumptions and Strategies

LEARNING OBJECTIVES

- Define and give examples of the basic assumptions underlying the analysis of behavior.
- List and describe the three levels of scientific understanding.
- Define and give examples of functional relations.
- Outline how experiments should be conducted to control for various threats to internal validity.
- Compare and contrast the role of variability in single-subject versus group research designs.
- List and define the essential components of experiments in applied behavior analysis research.
- Write a specific research question given a behavioral phenomenon of interest.
- Define and identify steady or stable rate responding.
- List and define the three elements of baseline logic.
- Discuss the value of establishing a steady baseline in applied behavior analysis.
- Outline guidelines for establishing a steady baseline in applied behavior analysis research.
- Identify four types of baseline data patterns.
- Discuss how single-subject research employs the inductive logic known as "affirmation of the consequent."
- Systematically manipulate independent variables and analyze their effects on treatment.

> Our technology of behavior change is also a technology of behavior measurement and of experimental design; it developed as that package, and as long as it stays in that package, it is a self-evaluating enterprise. Its successes are successes of known magnitude; its failures are almost immediately detected as failures; and whatever its outcomes, they are attributable to known inputs and procedures rather than to chance events or coincidences.
>
> —Donald M. Baer (personal communication, October 21, 1982)

Measurement shows whether, when, and how much behavior has changed, but measurement alone does not reveal why, or more accurately *how*, behavior change occurred. A useful technology of behavior change depends upon knowing the specific arrangements of environmental variables that produced behavior change. Without this knowledge, practitioners' efforts to change behavior would consist of procedures selected from a bag of tricks with little or no generality from one situation to the next.

The search for and demonstration of functional and reliable relations between socially important behavior and its controlling variables is a defining characteristic of applied behavior analysis. As articulated by Don Baer's chapter-opening statement, a major strength of applied behavior analysis is its insistence on experimentation as its method of proof, which enables and demands an ongoing, self-correcting search for effectiveness.

An experimental analysis must be accomplished to determine if and how a given behavior functions in relation to specific changes in the environment. This chapter introduces basic concepts and strategies that underlie the *analysis* in applied behavior analysis.[1] First, we briefly review some general conceptions of science and then continue with a discussion of two defining features and two assumptions about behavior that dictate the experimental methods most conducive to the subject matter. We identify the essential components of an experiment in applied behavior analysis and describe the basic logic that guides experimental methods in applied behavior analysis.

CONCEPTS AND ASSUMPTIONS UNDERLYING THE ANALYSIS OF BEHAVIOR

As discussed in Chapter 1, scientists share a set of common perspectives that include assumptions about the nature of the phenomena they study (determinism), the kind of information that should be gathered on the phenomena of interest (empiricism), the way questions about the workings of nature are most effectively examined (experimentation), and how an experiment's results should be judged (with parsimony and philosophic doubt). As Skinner (1953) noted, these attitudes apply to

all scientific disciplines: "The basic characteristics of science are not restricted to any particular subject matter" (p. 11).

The overall goal of science is complete understanding of the phenomena under study—socially significant behavior, in the case of applied behavior analysis. Scientists use a wide variety of methods that yield understanding at three levels: description, prediction, and control. First, systematic observation enables scientists to describe natural phenomena accurately. Descriptive knowledge of this type yields a collection of facts about the observed events—facts that can be quantified and classified, a necessary and important element of any scientific discipline.

A second level of scientific understanding is achieved when repeated observation shows that two events consistently covary. That is, the occurrence of one event (e.g., marriage) is associated with the occurrence of another event at some reliable degree of probability (e.g., longer life expectancy). The systematic covariation between two events—termed a *correlation*—can be used to predict the probability that one event will occur based on the presence of the other event.

The ability to predict successfully is a useful result of science; prediction allows preparation. However, the greatest potential benefits of science are derived from the third, and highest, level of scientific understanding, which comes from establishing experimental control. "The experimental method is a method for isolating the relevant variables within a pattern of events. Methods that depend merely on observed correlations, without experimental intervention, are inherently ambiguous" (Dinsmoor, 2003, p. 152).

Experimental Control: The Path to and Goal of Behavior Analysis

Behavior—the interaction between an organism and its environment—is best analyzed by measuring behavior change produced by imposed variations on the environment. This statement embodies the general strategy and the goal of behavior analytic research: to demonstrate that changes in the target behavior occurred *because of* experimentally manipulated changes in the environment.

Experimental control is achieved when a predictable change in behavior (the dependent variable) is reliably produced by the systematic manipulation of some aspect of the environment (the independent variable). Experimentally determining the effects of environmental manipulation on behavior and demonstrating that those effects can be reliably produced constitute the *analysis* in applied behavior analysis. An analysis of a behavior has been achieved when a reliable functional relation between the behavior and some specified aspect of the environment has been demonstrated convincingly. Knowledge of functional relations enables practitioners trained in applied behavior analysis to reliably alter behavior in meaningful ways.

An analysis of behavior "requires a believable demonstration of the events that can be responsible for the occurrence or nonoccurrence of that behavior. An experimenter has achieved an analysis of a behavior when he can exercise control over it" (Baer, Wolf, & Risley, 1968, p. 94).[2] That behavior analysts seek and value the experimental isolation of a given environmental variable of which a behavior is shown to be a function is often misinterpreted as support for a simplistic conception of behavioral causality. The fact that a behavior varies as a function of a given variable does not preclude that behavior also varying as a function of other variables. Thus, Baer and colleagues described an experimental analysis as a convincing demonstration that a variable *can* be responsible for the observed behavior change. Although a complete analysis (i.e., understanding) of a behavior requires an accounting of all of its multiple causes, an *applied* (i.e., technologically useful) analysis has been accomplished when the investigator has isolated an environmental variable (or group of variables that operate together as a treatment package) that reliably produces socially significant behavior change. An *applied* analysis also requires the target behavior to be a function of an environmental event that can be practically and ethically manipulated.

An experiment that shows convincingly that changes in the dependent variable are a function of the independent variable and not the result of uncontrolled or unknown variables is said to have a high degree of **internal validity**. A study with questionable internal validity yields no meaningful statements regarding functional relations between the variables examined in the experiment. Nor can a study without convincing internal validity be used as the basis for statements claiming generality of the findings to other persons, settings, and/or behaviors. **External validity** refers to the degree to which a study's results are generalizable to other subjects, settings, and/or behaviors. Strategies for assessing and extending the external validity of experimentally demonstrated functional relations are discussed in Chapter 10.

When planning an experiment and later when examining data from an ongoing study, investigators must always look out for threats to internal validity. Uncontrolled variables known or suspected to exert an influence on the dependent variable are called **confounding variables**. For example, suppose a researcher wants to analyze the effects of guided lecture notes on high school biology students' scores on next-day quizzes. Just one of many potential confounding variables the researcher would need to take into account would be each student's changing level of interest in and background knowledge about the specific curriculum content (e.g., a student's high score on a quiz following a lecture on sea life may be due to his prior knowledge about fishing, not the guided notes provided during that lecture).

A primary factor in evaluating the internal validity of an experiment is the extent to which it eliminates or controls the effects of confounding variables while still investigating the research questions of interest. Eliminating every possible source of uncontrolled variability in an experiment is an impossible ideal, but one all researchers should strive to attain. In reality, the goal of experimental design is eliminating as many uncontrolled variables as possible while holding constant the influence of all other variables except the independent variable, which is purposefully manipulated to determine its effects.

Behavior: Defining Features and Assumptions that Guide Its Analysis

Behavior is a difficult subject matter, not because it is inaccessible, but because it is extremely complex. Since it is a

> process, rather than a thing, it cannot easily be held still for observation. It is changing, fluid, evanescent, and for this reason it makes great technical demands upon the ingenuity and energy of the scientist.
>
> —B. F. Skinner (1953, p. 15)

How a scientific discipline defines its subject matter exerts profound influence and imposes certain constraints on the experimental strategies that will be most effective in an understanding of it. Experimental methods in behavior analysis are guided by two defining features of behavior: (a) Behavior is an individual phenomenon and (b) behavior is dynamic and continuous; and two assumptions concerning its nature: (a) behavior is determined and (b) behavioral variability is extrinsic to the organism.

Behavior Is an Individual Phenomenon

If behavior is defined as a person's interaction with the environment, it follows that a science seeking to discover general principles or laws that govern behavior must study the behavior of individuals. Groups of people do not behave; individual people do. Thus, the experimental strategy of behavior analysis uses within-subject (or single-case) methods of analysis.

The average performance of a group of individuals often yields interesting and useful information and may, depending on the methods by which individuals were selected to be in the group, enable probability statements about the average performance within the larger population represented by the group. However, "group data" provide no information about the behavior of any individual or how any individual might perform in the future. For example, although administrators and taxpayers may be justifiably interested in the average increase in students' reading comprehension from grade level to grade level, such information is of little use to the classroom teacher who must decide how to improve a given student's comprehension skills.

Nonetheless, learning how behavior–environment relations work with many individuals is vital. A science of behavior contributes to a useful technology of behavior change only to the extent that it discovers functional relations with generality across individuals. The issue is how to achieve that generality. Behavior analysts have found that discovery of behavioral principles with generality across persons is best accomplished by replicating demonstrated functional relations with additional subjects.

Behavior Is Dynamic and Continuous

Just as behavior cannot take place in an environmental void, so must behavior occur in time. "Behavior is a continuous process, from birth through to death. The problem is that this stream of change cannot be stopped for scientific purposes" (Keenan & Dillenburger, 2000, p. 26). Measuring behavior as it changes over time is a fundamental feature of applied behavior analysis. Therefore, single measures, or even multiple measures sporadically dispersed over time, cannot provide an adequate description of behavior change. Only continuous measurement yields a complete record of behavior as it occurs in context with its environmental influences. Because true continuous measurement is seldom feasible in applied settings, systematic repeated measurement of behavior (as described in Chapters 4 and 5) is the hallmark of applied behavior analysis.

Behavior Is Determined

As discussed in Chapter 1, all scientists assume that the universe is a lawful and orderly place and that natural phenomena occur in relation to other natural events.

> The touchstone of all scientific research is order. In the experimental analysis of behavior, the orderliness of relations between environmental variables and the subject's behavior is at once the operating assumption upon which the experimenter proceeds, the observed fact that permits doing so, and the goal that continuously focuses experimental decisions. (Johnston & Pennypacker, 1993a, p. 238)

Behavior analysts consider behavior to be a natural phenomenon that, like all natural phenomena, is determined. Although determinism must always remain an assumption—it cannot be proven—it is an assumption with strong empirical support.

> Data gathered from all scientific fields indicate that *determinism* holds throughout nature. It has become clear that the *law of determinism*, that is, that all things are determined, holds for the behavioral area also. . . . When looking at actual behavior we've found that in situation 1, behavior is caused; in situation 2, behavior is caused; in situation 3, behavior is caused; . . . and in situation 1001, behavior is caused. Every time an experimenter introduces an independent variable that produces some behavior or some change in behavior, we have further *empirical* evidence that behavior is caused or deterministic. (Malott, General, & Snapper, 1973, pp. 170, 175)

Behavioral Variability Is Extrinsic to the Organism

When repeated measures reveal inconsistent responding by a subject when environmental conditions are held constant, the data are said to display variability. The experimental approach most commonly used in psychology and other social and behavioral sciences (e.g., education, sociology, political science) makes two assumptions about such variability: (a) Behavioral variability is an intrinsic characteristic of the organism, and (b) behavioral variability is distributed randomly among individuals in any given population. These assumptions have critical methodological implications: (a) Attempting to control or investigate variability is a waste of time—it simply exists, it's a given; and (b) by averaging the performance of individual subjects within large groups, the random nature of variability can be statistically controlled or canceled out. Both assumptions about behavioral variability are likely false (empirical evidence points in the opposite direction), and the methods they encourage are detrimental to a science of behavior. "Variables are not canceled statistically. They are simply buried so their effects are not seen" (Sidman, 1960, p. 162).[3]

Behavior analysts approach variability in their data quite differently. A fundamental assumption underlying the design and guiding the conduct of experiments in behavior analysis is that, rather than being an intrinsic characteristic of the organism, behavioral variability is the result of environmental influence: the

independent variable with which the investigator seeks to produce change, an uncontrolled aspect of the experiment itself, and/or an uncontrolled or unknown factor outside the experiment.

The assumption of extrinsic variability yields the following methodological implication: Instead of averaging the performance of many subjects in an attempt to mask variability (and forfeiting the opportunity to understand and control it), behavior analysts experimentally manipulate factors suspected of causing the variability. Searching for the causal factors contributes to the scientific understanding of behavior, because experimentally demonstrating a source of variability suggests another functional relation. In some studies, "tracking down these answers may even turn out to be more rewarding than answering the original experimental question" (Johnston & Pennypacker, 1980, p. 226).

From a purely scientific viewpoint, experimentally tracking down sources of variability is always the preferred approach. However, the applied behavior analyst, with a problem to solve, must often take variability as it presents itself (Sidman, 1960). Sometimes the applied researcher has neither the time nor the resources to experimentally manipulate even suspected and likely sources of variability (e.g., a teacher who interacts with a student for only part of the day has no hope of controlling the many variables outside the classroom that influence the student's behavior). In most settings, applied behavior analysts seek treatment variables robust enough to overcome variability induced by uncontrolled factors (Baer, 1977b).

COMPONENTS OF APPLIED BEHAVIOR ANALYSIS EXPERIMENTS

> Nature to be commanded must be obeyed. . . . But, that coin has another face. Once obeyed, nature can be commanded.
>
> —B. F. Skinner (1956, p. 232)

Experimentation is the scientist's way of discovering nature's rules. Discoveries that prove valid and reliable can contribute to a technology of effective behavior change. Experiments in applied behavior analysis include these essential components:

- Research question
- At least one participant (subject)
- At least one behavior (dependent variable)
- At least one setting
- A system for measuring the behavior and ongoing visual analysis of the data
- At least one intervention or treatment condition (independent variable)
- Manipulation of the independent variable to determine its effects on the dependent variable (experimental design)

Research Question

> We conduct experiments to find out something we do not know.
>
> —Murray Sidman (1960, p. 214)

For the applied behavior analyst, Sidman's "something we do not know" is cast in the form of a question about the existence and/or specific nature of a functional relation between meaningful improvement in socially significant behavior and one or more of its controlling variables. A **research question** specifies what the investigator wants the experiment to answer. Research questions in applied behavior analysis take many forms; most questions are one of four types:

- ***Demonstration***—to what extent will this intervention work (i.e., change the behavior of interest)?
- ***Parametric***—does more or less of the intervention work better?
- ***Component***—how effective is the intervention when various components are added or subtracted?
- ***Comparative***—does one intervention work better than another? (adapted from Wolery, Lane, & Common, 2018)

In some published reports of applied behavior analysis studies, the investigators state their research question(s) explicitly, as in these examples:

- The research questions were as follows: 1. What are the effects of a self-monitoring package with conferencing on the completion of spelling and math homework for students with disabilities in an inclusive general education classroom? 2. What are the effects of a self-monitoring package with conferencing on the accuracy of spelling and math homework for students with disabilities in an inclusive general education classroom? 3. If gains are made in completion and accuracy of spelling and math homework, will students maintain them 2 and 3 weeks after the intervention has ended? (Falkenberg & Barbetta, 2013, pp. 191–192)
- Four primary research questions were addressed: What are the effects of training middle school students with learning disabilities to recruit teacher attention in the special education classroom on (a) the number of recruiting responses they emit in the general education classroom, (b) the number of teacher praise statements received by the students in the general education classroom, (c) the number of instructional feedback statements received by the students in the general education classroom, and (d) the students' academic productivity and accuracy in the general education classroom? (Alber, Heward, & Hippler, 1999, p. 255)
- Experiment 1—What are the differences in responding under the "no programmed consequence," "praise," and "known reinforcer" conditions? To what extent will 3 to 5 "pairing" sessions per day affect responding during "praise" sessions conducted at the beginning and end of each day? (Axe & Laprime, 2017, p. 330)

More often, however, the research question is implicit within a statement of the study's purpose. For example:

- [T]he purpose of this study was to determine the effects of self-graphing on the quantity and quality of writing for students with high-incidence disabilities. (Stolz, Itoi, Konrad, & Alber-Morgan, 2008, p. 174)
- [T]he purpose of this study was to evaluate the extent to which lag schedules of reinforcement produced responses

that were already in the existing repertoires of three children diagnosed with autism, as well as the extent to which lag schedules produced novel responses. (Contreras & Betz, 2016, p. 5)

- [T]he purpose of this study was to examine the use of habit reversal to decrease filled pauses that occur during public speaking. (Mancuso & Miltenberger, 2016, p. 188)
- [W]e sought to determine if teaching three children with autism to monitor their peers' reading responses would lead to the acquisition of sight words. (Taylor, DeQuinzio, & Stine, 2012, p. 815)
- The purpose of this study was to determine whether naturally occurring meals would affect performance adversely during postmeal sessions in which highly preferred food was used as reinforcement. (Zhou, Iwata, & Shore, 2002, pp. 411–412)

Whether the research question is stated explicitly in the form of a question or implicit within a statement of purpose, all aspects of the experiment's design and conduct should follow from it.

> A good design is one that answers the question convincingly, and as such needs to be constructed in reaction to the question and then tested through arguments in that context (sometimes called "thinking through"), rather than imitated from a textbook. (Baer, Wolf, & Risley, 1987, p. 319)

Participant

Experiments in applied behavior analysis are most often referred to as *single-subject* or **single-case designs**. These terms do not mean that applied behavior analysis studies are conducted with only one subject (though some are); they denote an experimental logic for detecting and analyzing functional relations between environmental variables and behavior change that employs each subject as her own control.[4] Repeated measures of each subject's behavior are obtained as she is exposed to each condition of the study (e.g., the presence and absence of the independent variable). A subject is often exposed to each condition several times over the course of an experiment. Measures of the subject's behavior during each phase of the study provide the basis for comparing the effects of experimental variables as they are presented or withdrawn in subsequent conditions.

Although most applied behavior analysis studies involve more than one subject, each subject's data are graphed and analyzed separately.[5] Instead of using *single-case* or *single-subject design* to refer to an experiment in which each subject serves as his or her own control, some authors use more aptly descriptive terms such as *within-subject design*, *intrasubject design*, and *repeated-measures design*.[6]

Sometimes the applied behavior analyst is interested in assessing the overall effect of a treatment variable within a group of subjects—for example, the number of homework assignments completed by members of a class of fifth-grade students. In such a case, the total number of assignments completed may be measured, graphed, and analyzed as a dependent variable. However, unless each student's data are individually graphed and interpreted, no individual student's behavior has been analyzed, and the data for the group may not be representative of any individual subject.

Use of a single participant, or a small number of participants, each of whom is considered an intact experiment, stands in sharp contrast to the between-group designs traditionally used in psychology and the other social sciences that employ large numbers of subjects. Proponents of between-group designs contend that using a large number of subjects controls for the variability discussed earlier and increases the generality (or external validity) of any findings to the population from which the subjects were selected. The advantages and disadvantages of an experimental approach based on within-subject comparisons of the behavior of individual subjects versus comparisons of the average performance of different groups of subjects is discussed in Chapter 10.

Behavior: Dependent Variable

The target behavior in an applied behavior analysis experiment, or more precisely a measurable dimensional quantity of that behavior (e.g., rate, duration), is called the **dependent variable**. It is so labeled because the experiment is designed precisely to determine whether the behavior is, in fact, *dependent on* (i.e., a function of) the independent variable(s) manipulated by the investigator. (Criteria and procedures for selecting and defining response classes that meet the *applied* requirements as dependent variables in applied behavior analysis are described in Chapter 3.)

In some studies multiple behaviors are measured. One reason for measuring multiple behaviors is to provide data patterns that can serve as controls for evaluating and replicating the effects of an independent variable as it is applied sequentially to each of the behaviors.[7] A second reason for multiple dependent measures is to assess the presence and extent of the independent variable's effects on behaviors other than the response class to which it was applied directly. This strategy can reveal whether the independent variable has collateral effects—either desired or undesired—on other behaviors of interest. Such behaviors are referred to as secondary dependent variables. The experimenter obtains regular measures of their rate of occurrence, though perhaps not as often as measures of the primary dependent variable are recorded.

Still another reason for measuring multiple behaviors is to determine whether changes in the behavior of a person other than the subject occur during the course of an experiment and whether such changes might in turn explain observed changes in the subject's behavior. This strategy is implemented primarily to assess the effects of a suspected confounding variable: The additional behavior or behaviors measured are not true dependent variables in the sense of undergoing analysis. For example, in a classic study analyzing the effects of the self-recording by a junior high school girl on her classroom study behavior, Broden, Hall, and Mitts (1971) observed and recorded the number of times the girl's teacher paid attention to her throughout the experiment. If the frequency of teacher attention had covaried with changes in the girl's study behavior, a functional relation between self-recording and study behavior would not have been demonstrated. In that case, teacher attention would likely have been identified as a potential confounding variable,

and the focus of the investigation shifted to include efforts to experimentally control it (i.e., to hold teacher attention constant) or to systematically manipulate and analyze its effects. However, the data revealed no correlation between teacher attention and study behavior during the first four phases of the experiment, when concern was highest that teacher attention may have been a confounding variable.

Setting

> Control the environment and you will see order in behavior.
> —B. F. Skinner (1967, p. 399)

Functional relations are demonstrated when observed variations in behavior can be attributed to specific operations imposed on the environment. *Experimental control* is achieved when the systematic manipulation of some aspect of the subject's environment (the independent variable) reliably produces a predictable change in behavior (the dependent variable). To make such an attribution, the investigator must, among other things, control two sets of environmental variables. First, the investigator must control the presence, absence, or value of the independent variable in every phase of the experiment. Second, the investigator must control, by holding constant, all other aspects of the experimental setting—**extraneous variables**—to prevent unplanned environmental variation. These two operations—precisely manipulating the independent variable and maintaining the constancy of every other relevant aspect of the experimental setting—define the second meaning of *experimental control.*

In basic laboratory research, experimental space is designed and furnished to maximize experimental control. Lighting, temperature, and sound, for example, are all held constant, and programmed apparatus virtually guarantee the presentation of antecedent stimuli and the delivery of consequences as planned. Applied behavior analysts carry out research where socially important behaviors naturally occur—classrooms, homes, workplaces, and community settings. It is impossible to control every feature of an applied environment, and to add to the difficulty, subjects are typically in the experimental setting for a small portion of each day, bringing with them the influences of events and contingencies operating in other settings.

In spite of the complexity and ever-changing nature of applied settings, the behavior analyst must make every effort to hold constant all seemingly relevant aspects of the environment. When unplanned variations are noted, the investigator must wait out their effects or try to incorporate them into the design of the experiment. In any event, repeated measures of the subject's behavior provide the barometer for assessing whether unplanned environmental changes are of concern.

Applied studies are often conducted in more than one setting. Researchers sometimes use concurrent measures of the same behavior obtained across settings as controls for analyzing the effects of an independent variable that is sequentially applied to the behavior in each setting.[8] Data are also collected in multiple settings to assess the extent to which behavior changes observed in the primary setting have occurred in the other setting(s).

Measurement System and Ongoing Visual Analysis

Beginning students of behavior analysis sometimes believe the discipline is preoccupied with issues and procedures related to the observation and measurement of behavior. They want to get on with the analysis. Measuring behavior change over time is the principal feature of single-case experimental method. The results of any experiment can be presented and interpreted only in terms of what was measured, and the observation and recording procedures used in the study determine not only what was measured but also how well it was measured (i.e., how representative of the subject's actual behavior is the estimate provided by the experimental data—all measurements of behavior, no matter how frequent and technically precise, are estimates of true values).

Observation and recording procedures must be conducted in a standardized manner throughout each session of an experiment. Standardization involves every aspect of the measurement system, from the definition of the target behavior (dependent variable) to the scheduling of observations to the manner in which the raw data are transposed from recording sheets to session summary sheets to the way the data are graphed. As detailed in Chapter 5, unintended or imposed changes in measurement tactics can result in unwanted variability or confounded treatment effects.

The previous chapter outlined advantages that accrue to the behavioral researchers and practitioners who maintain direct contact with the experimental data by ongoing visual inspection of graphic displays. Behavior analysts must recognize changes in level, trend, and degree of variability as these changes develop in the data. Because behavior is a continuous, dynamic phenomenon, experiments designed to discover its controlling variables must enable the investigator to inspect and respond to the data continuously as the study progresses. Only in this way can the behavior analyst be ready to manipulate features of the environment at the time and in the manner that will best reveal functional relations and minimize the effects of confounding variables.

Intervention or Treatment: Independent Variable

Behavior analysts seek functional relations between behavior change and environmental variables. The particular aspect of the environment that the experimenter manipulates to find out whether it affects the subject's behavior is called the **independent variable**. Sometimes called the *experimental variable*, *intervention*, or *treatment*, this component of an experiment is called the independent variable because the researcher controls or manipulates it independent of the subject's behavior or any other event. (As we will soon see, though, manipulating the independent variable without regard to measures of the dependent variable is unwise.) Whereas changes in the experimental setting required to conduct the study (e.g., the presence of observers to measure behavior) are made with the goal of minimizing their effects on the dependent variable, "changes in the independent variable are arranged by the experimenter in order to maximize . . . its influence on responding" (Johnston & Pennypacker, 1980, p. 260).

Manipulation of the Independent Variable: Experimental Design

Experimental design refers to the particular arrangement of conditions in a study so that meaningful comparisons of the effects of the presence, absence, or different values of the independent variable can be made. Independent variables can be introduced, withdrawn, increased or decreased in value, or combined across behaviors, settings, and/or subjects in an infinite number of ways.[9] However, with respect to the behavior of a given subject, at a given point in time in a given setting, an experimenter can make only two basic kinds of independent variable changes: introduce a new condition or reinstate an old condition. In essence, "Experimental designs are merely temporal arrangements of various new and old conditions across behaviors and settings in ways that produce data that are convincing to the investigator and the audience" (Johnston & Pennypacker, 1980, p. 270).

In the simplest case—from an analytic perspective, but not necessarily a practical point of view—an independent variable can be either present or absent during each time period or phase of a study. When the independent variable is in either of these conditions during a study, the experiment is termed a *nonparametric study*. In contrast, a **parametric analysis** seeks to discover the differential effects of a range of values of the independent variable. Lerman, Kelley, Vorndran, Kuhn, and LaRue (2002) conducted a parametric analysis when they assessed the effects of different reinforcer magnitudes (20, 60, or 300 sec of access to toys or escape from demands) on the duration of postreinforcement pause and resistance to extinction. Examples of different values of independent variables in other parametric studies: short (2 sec), progressive (2 sec to 20 sec), and long (20 sec) intertrial-interval durations during language training (Cariveau, Kodak, & Campbell, 2016); and teacher praise rates of 1, 2, or 8 per minute during small-group instruction (Kranak, Alber-Morgan, & Sawyer, 2017).

When an investigator is interested in comparing the effects of two or more treatment alternatives, multiple independent variables become part of the experiment. For example, perhaps two separate treatments are evaluated as well as the effects of a third treatment, which is a combination of both variables. However, even in studies with multiple independent variables, the researcher must heed a simple but fundamental rule of experimentation: *Change only one variable at a time*. Failure to heed this rule eliminates the possibility of attributing any measured behavior change to a specific independent variable.

If two or more variables are altered simultaneously and changes in the dependent variable are noted, no conclusions can be made with regard to the contribution of any one of the altered variables to the behavior change. If two variables changed together, both could have contributed equally to the resultant behavior change; one variable could have been solely, or mostly, responsible for the change; or one variable may have had a negative or counterproductive effect, but the other independent variable was sufficiently strong to overcome this effect, resulting in a net gain. Any of these explanations or combinations may have accounted for the change.

As stated previously, applied behavior analysts often conduct their experiments in "noisy" environments where effective treatments are required for reasons related to personal safety or exigent circumstances. In such cases, applied behavior analysts sometimes "package" multiple and well-documented and effective treatments, knowing that multiple independent variables are being introduced. As implied earlier, a package intervention is one in which multiple independent variables are being combined or bundled into one program (e.g., token reinforncement + praise + self-recording + time-out). However, from the perspective of experimental analysis, the rule to change only one variable at a time still holds. When manipulating a treatment package, the experimenter must ensure that the entire package is presented or withdrawn each time a manipulation occurs. In this situation, it is important to understand that the entire package is being evaluated, not the discrete components that make up the package. If, at a later time, the analyst wishes to determine the relative contributions of each part of the package, she would need to carry out a component analysis. Chapters 8 and 9 describe experimental tactics for component analyses.

No off-the-shelf experimental designs are available for a given research problem. Investigators must avoid following textbook "designs" that (a) require a priori assumptions about the nature of the functional relation one seeks to investigate and (b) may be insensitive to unanticipated changes in behavior. Researchers should select and combine experimental tactics that best fit their research questions, while standing ever ready to "explore relevant variables by manipulating them in an improvised and rapidly changing design" (Skinner, 1966, p. 21).

Simultaneous with, and to a large degree responsible for, the growth and success of applied behavior analysis have been the development and refinement of a powerful group of extremely flexible experimental tactics for analyzing behavior–environment relations. We describe the most widely used of these tactics in Chapters 8 and 9. However, to effectively select, modify, and combine these tactics into convincing experiments, behavior analysts must fully understand the experimental reasoning, or logic, that provides the foundation for within-subject experimental comparisons.

STEADY STATE STRATEGY AND BASELINE LOGIC

Steady state responding—"a pattern of responding that exhibits relatively little variation in its measured dimensional quantities over a period of time" (Johnston & Pennypacker, 1993a, p. 199)—provides the basis for a powerful form of experimental reasoning called baseline logic. **Baseline logic** entails three elements—prediction, verification, and replication—each of which depends on an overall experimental approach called steady state strategy. **Steady state strategy** entails exposing a subject to a given condition while trying to eliminate or control any extraneous influences on the behavior and obtaining a stable pattern of responding before introducing the next condition.

Nature and Function of Baseline Data

Behavior analysts discover behavior–environment relations by comparing data from repeated measures of a subject's behavior under the different conditions of the experiment. The most

common method of evaluating the effects of a given variable is to impose it on an ongoing measure of behavior obtained in its absence. These original data serve as the **baseline** against which any observed changes in behavior when the independent variable is introduced can be compared. A baseline serves as a control condition; baseline does not necessarily mean the absence of instruction or treatment, only the absence of the independent variable of experimental interest.

Why Establish a Baseline?

From a purely scientific or analytic perspective, the primary purpose for establishing a baseline level of responding is to use the subject's performance in the absence of the independent variable as an objective basis for detecting the effects of the independent variable. Obtaining baseline data can yield several applied benefits. For one, systematic observation of the target behavior before a treatment variable is introduced provides the opportunity to look for and note environmental events that occur just before and just after the target behavior. Such empirically obtained descriptions of antecedent-behavior-consequent correlations are often invaluable in planning an effective intervention (see Chapter 25). For example, baseline observations revealing that a child's disruptive outbursts are consistently followed by parent or teacher attention can be used in designing an intervention of ignoring outbursts and contingent attention following desired behavior.

Second, baseline data can provide valuable guidance in setting initial criteria for reinforcement, a particularly important step when a contingency is first put into effect (see Chapter 11). If the criteria are too high, the subject never comes into contact with the contingency; if they are too low, little or no improvement can be expected.

From a practical perspective, a third reason for collecting baseline data concerns the merits of objective measurement versus subjective opinion. The results of systematic baseline measurement may convince the behavior analyst or significant others to reconsider the necessity and value of attempting to change the behavior. For example, a behavior targeted for intervention because of several recent and extreme instances is no longer targeted because baseline data show it is decreasing. Or, perhaps a behavior's topography attracted undue attention from teachers or parents, but objective baseline measurement over several days reveals the behavior is not occurring at a rate or level that warrants intervention.

Baseline Data Patterns

Figure 7.1 shows examples of four data patterns sometimes generated by baseline measurement. These hypothetical baselines represent only four examples of the wide variety of baseline data patterns experimenters and practitioners encounter. The potential combinations of different levels, trends, and degrees of variability are infinite. Nevertheless, in an effort to provide guidance to the beginning behavior analyst, some general statements will be given about the experimental decisions that might be warranted by the data patterns shown in Figure 7.1.

Graph A shows a relatively **stable baseline**. The data present no evidence of an upward or downward trend, and all of the measures fall within a small range of values. A stable baseline provides the most desirable basis, or context, against which to look for effects of an independent variable. If changes in level, trend, and/or variability coincide with the introduction of an independent variable on a baseline as stable as that shown in Graph A, one can reasonably suspect that those changes may be related to the independent variable.

The data in Graphs B and C represent an **ascending baseline** and a **descending baseline**, respectively. The data path in Graph B shows an increasing trend in the behavior over

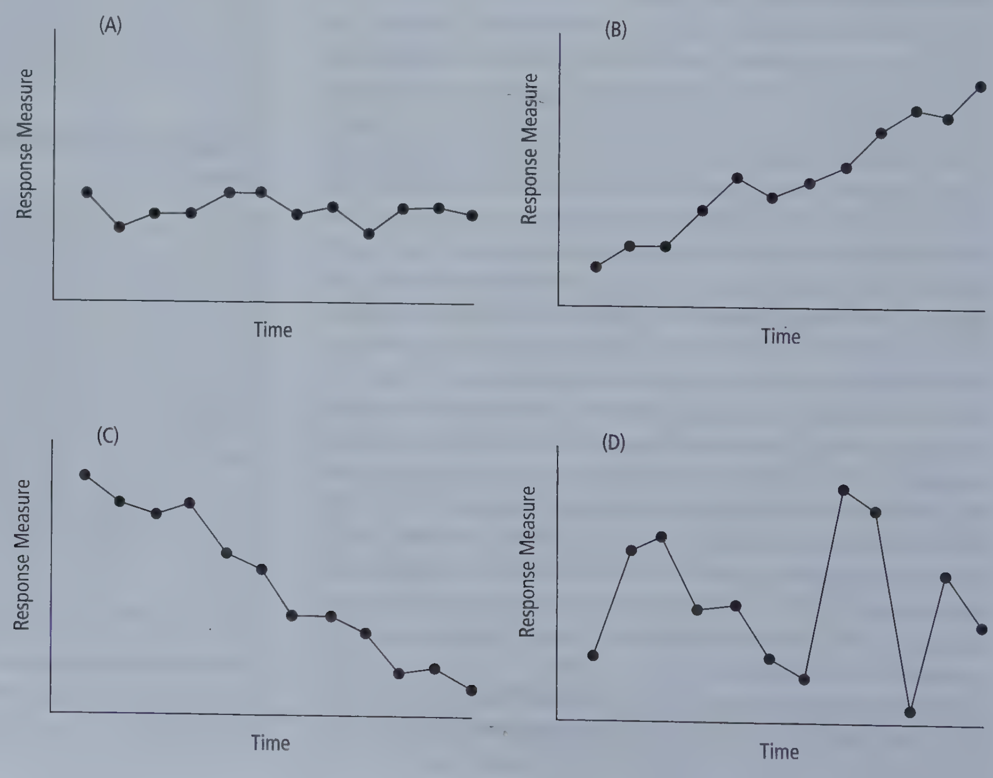


Figure 7.1 Data patterns illustrating stable (A), ascending (B), descending (C), and variable (D) baselines.

time, whereas the data path in Graph C presents a decreasing trend. The applied behavior analyst must treat ascending and descending baseline data cautiously. Applied behavior analysts select dependent variables because they represent target behaviors that need to be changed. But ascending and descending baselines reveal behaviors in the process of changing. The effects of an independent variable introduced at this point are likely to be obscured or confounded by the variables responsible for the already-occurring change. But what if the applied investigator needs to change the behavior immediately? The applied perspective can help solve the dilemma.

Whether a treatment variable should be introduced depends on whether the trending baseline data represent improving or deteriorating performance. When an ascending or a descending baseline represents behavior change in the therapeutically desired direction, the investigator should withhold treatment and continue to monitor the dependent variable under baseline conditions. When the behavior ceases to improve (as evidenced by stable responding) or begins to deteriorate, the independent variable can be applied. If the trend does not level off and the behavior continues to improve, the original problem may no longer be present, leaving no reason for introducing the treatment as planned (although the investigator might be motivated to isolate and analyze the variables responsible for the "spontaneous" improvement). Introducing an independent variable to an already-improving behavior makes it difficult, and often impossible, to claim any continued improvement as a function of the independent variable.

An ascending or descending baseline that represents significantly deteriorating performance signals an immediate application of the independent variable. From an applied perspective, the decision to intervene is obvious: The subject's behavior is worsening, and a treatment designed to improve it should be introduced. An independent variable capable of effecting desired behavior change in spite of other variables "pushing" the behavior in the opposite direction is most likely a robust variable, one that will be a welcome addition to the behavior analyst's list of effective treatments. The decision to introduce a treatment variable on a descending baseline is also a sound one from an analytic perspective, which will be discussed in the next section.

Graph D in Figure 7.1 shows a highly unstable or **variable baseline.** The data in Graph D show just one of many possible patterns of unstable responding. The data points do not consistently fall within a narrow range of values, nor do they suggest any clear trend. Introducing the independent variable in the presence of such variability is unwise from an experimental standpoint. Variability is assumed to be the result of environmental variables, which in the case shown by Graph D seem to be operating in an uncontrolled fashion. Before the researcher can analyze the effects of an independent variable effectively, these uncontrolled sources of variability must be isolated and controlled.

Stable baseline responding provides an index of the degree of experimental control the researcher has established. "If sufficiently stable responding cannot be obtained, the experimenter is in no position to add an independent variable of suspected but unknown influence. To do so would be to compound confusion and lead to further ignorance" (Johnston & Pennypacker, 1980, p. 229).

Again, however, applied considerations must be balanced against purely scientific pursuits. The applied problem may be one that cannot wait to be solved (e.g., severe self-injurious behavior). Or, confounding variables in the subject's environment and the setting(s) of the investigation may simply be beyond the experimenter's control.[10] In such situations, the independent variable is introduced with the hope of producing stable responding in its presence. Sidman (1960) agreed that "the behavioral engineer must ordinarily take variability as he finds it, and deal with it as an unavoidable fact of life" (p. 192).

Prediction

Prediction can be defined as "the anticipated outcome of a presently unknown or future measurement. It is the most elegant use of quantification upon which validation of all scientific and technological activity rests" (Johnston & Pennypacker, 1980, p. 120). Figure 7.2 shows a series of hypothetical measures representing a stable pattern of baseline responding. The consistency of the first five data points in the series encourages the prediction that—if no changes occur in the subject's environment—subsequent measures will fall within the range of values obtained thus far. Indeed, a sixth measure is taken that gives credence to this prediction. The same prediction is made again, this time with more confidence, and another measure of behavior shows it to be correct. Throughout a baseline (or any other experimental condition), predictions are made and confirmed until the investigator has every reason to believe that the response measure will not change appreciably under the present conditions. Open data points in Figure 7.2 represent predicted measures of future responding. Given the stability of the obtained measures, few experienced scientists would quarrel with the prediction.

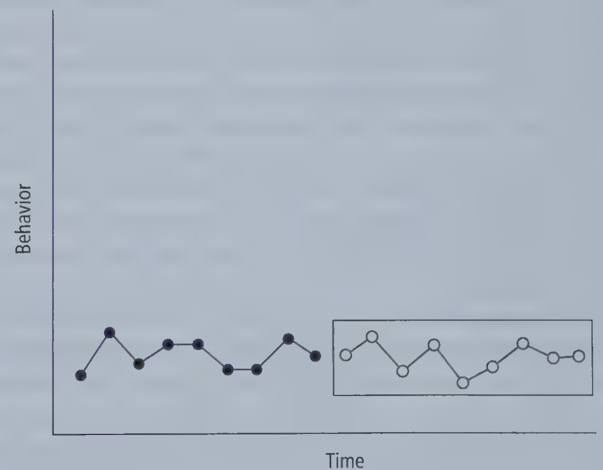


Figure 7.2 Solid data points represent actual measures of behavior that might be generated in a stable baseline; open data points within the box represent the level of responding predicted on the basis of the obtained measures, should the environmental conditions remain unchanged.

How many measures must be taken before an experimenter or practitioner can predict future behavior with confidence? There are no criteria or pre-determined number of data points that ensure confident prediction. However, the following statements can serve as guidelines.

- Baer and colleagues (1968) recommended continuing measurement until "stability is clear."
- All things being equal, many measures are better than a few; and the longer the period of time in which stable responding is obtained, the better the predictive power of those measures.
- If the experimenter is not sure whether measurement has revealed stable responding, in all likelihood it has not, and more data should be collected before the independent variable is introduced.
- The investigator's knowledge of the characteristics of the behavior being studied under constant conditions is invaluable in deciding when to terminate baseline measurement and introduce the independent variable. That knowledge can be drawn from personal experience in obtaining stable baselines on similar response classes and from familiarity with patterns of baseline responding found in the published literature.

Statements such as "collect baseline data for at least five sessions" or "obtain baseline measures for 2 consecutive weeks" are misguided or naive. Depending on the situation, five baseline data points or 2 weeks of baseline data may or may not provide a convincing picture of steady state responding. The question that must be addressed is this: Are the data sufficiently stable to provide a basis for comparison to measures collected from other conditions? This question can be answered only by ongoing prediction and confirmation with repeated measures obtained in an environment where all relevant conditions are held constant.

Behavior analysts are often interested in analyzing functional relations between an instructional variable and the acquisition of new skills. In such situations it is sometimes assumed that baseline measures are zero. For example, one would expect repeated observations of a child who has never tied her shoes to yield a perfectly stable baseline of zero correct responses. However, casual observations that have never shown a child to use a particular skill do not constitute a scientifically valid baseline and should not be used to justify any claims about the effects of instruction. It could be that, if given repeated opportunities to respond, the child would begin to emit the target behavior at a nonzero level.

Practice effects refer to improvements in performance resulting from repeated opportunities to emit the behavior so that baseline measurements can be obtained. For example, attempting to obtain stable baseline data for students answering arithmetic problems may result in improved performance simply because of the repeated practice inherent in the measurement process. Practice effects can confound a study, making it impossible to separate and account for the effects of practice and instruction on the student's final performance. Repeated baseline measures should be used either to reveal the existence or to demonstrate the nonexistence of practice effects. When practice effects are suspected or found, baseline data collection should be continued until steady state responding is attained.

The necessity to demonstrate a stable baseline and to control for practice effects empirically does not require applied behavior analysts to withhold needed treatment or intervention. Nothing is gained by collecting unduly long baselines of behaviors that cannot reasonably be expected to be in the subject's repertoire. Many behaviors cannot be emitted unless the subject is competent in certain prerequisite behaviors. For example, there is no legitimate possibility of a child's tying his shoes if he currently does not pick up the laces, or of a student's solving division problems if she cannot subtract and multiply. Obtaining extended baseline data in such cases is unnecessary pro forma measurement. Such measures do "not so much represent zero behavior as zero opportunity for behavior to occur, and there is no need to document at the level of well-measured data that behavior does not occur when it cannot" (Horner & Baer, 1978, p. 190).

Fortunately, applied behavior analysts need neither abandon the use of steady state strategy nor repeatedly measure nonexistent behavior at the expense of a treatment condition. The multiple probe design, described in Chapter 9, embodies steady state logic to analyze functional relations between instruction and the acquisition of behaviors shown to be nonexistent in the subject's repertoire prior to the introduction of the independent variable.

Affirmation of the Consequent

The predictive power of steady state responding enables the behavior analyst to employ a form of inductive logic known as **affirmation of the consequent** (Johnston & Pennypacker, 1980). The reasoning behind affirmation of the consequent begins with a true antecedent–consequent (if-A-then-B) statement and proceeds as follows:

1. If A is true, then B is true.
2. B is found to be true.
3. Therefore, A is true.

The behavior analyst's version goes like this:

1. If the independent variable is a controlling factor for the behavior (A), then the data obtained in the presence of the independent variable will show that the behavior has changed (B).
2. When the independent variable is present, the data show that the behavior has changed (B is true).
3. Therefore, the independent variable is a controlling variable for the behavior (therefore, A is true).

Although the logic is flawed—other factors could be responsible for the truthfulness of A—a successful (i.e., convincing) experiment affirms several if-A-then-B possibilities, each one reducing the likelihood that factors other than the independent variable are responsible for the observed changes in behavior.

Figures 7.3 to 7.5 illustrate how prediction, verification, and replication are employed in a hypothetical experiment using a reversal design, one of the more powerful analytic tactics used by behavior analysts (see Chapter 8). The data in Figure 7.3

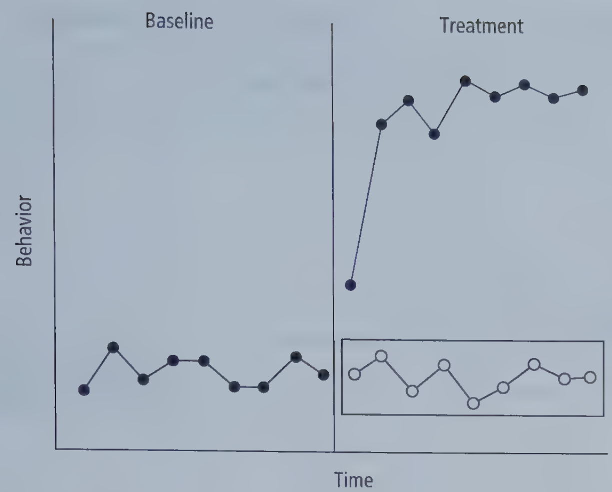


Figure 7.3 Affirmation of the consequent supporting the possibility of a functional relation between the behavior and treatment variable. The measures obtained in the presence of the treatment variable differ from the predicted level of responding in the absence of the treatment variable (open data points within the boxed area). Solid data points represent an A-B design.

show a successful affirmation of the consequent. Steady state responding during Baseline enabled a prediction that, if no changes were made in the environment, continued measurement would yield data similar to those represented by the open-circle data points within the boxed area. The experimenter also predicted that (more precisely, questioned whether) the treatment condition would produce a change in the behavior. The independent variable was introduced, and repeated measures of the dependent variable showed that the behavior did indeed change (solid data points in Treatment condition).

These data enable two comparisons, one real and one hypothetical. First, the difference between the obtained measures in the presence of the independent variable and those obtained during baseline supports the prediction that treatment would change the behavior and represents the extent of a possible effect of the independent variable. The second comparison is between the data obtained in the treatment condition and the predicted measures had the treatment variable not been introduced (i.e., the open data points within the boxed area of Figure 7.3). This comparison represents the behavior analyst's hypothetical approximation of the ideal, but impossible-to-achieve, experimental design: the simultaneous measurement and comparison of the behavior of an individual subject in both the presence and absence of the treatment variable (Risley, 1969).

Although the data in Figure 7.3 affirm the initial antecedent–consequent statement—a change in the behavior was observed in the presence of the independent variable—asserting a functional relation between the treatment variable and behavior change is unwarranted. The experimenter has not yet ruled out the possibility of other variables being responsible for the behavior change. Perhaps some factor within or outside the experimental setting occurred at the same time the independent variable was introduced, and that factor is responsible for the change in behavior.

Although two-phase experiments consisting of a pretreatment baseline condition followed by a treatment condition (i.e., an **A-B design**) enable neither verification of the prediction of continued responding at baseline levels nor replication of the effects of the independent variable, studies using A-B designs can nevertheless contribute important and useful findings to research and practice (e.g., Azrin & Wesolowski, 1974; Ghaemmaghami, Hanley, & Jessel, 2016; Krueger, Rapp, Ott, Lood, & Novotny, 2013; Reid, Parsons, Phillips, & Green, 1993).

A firmer statement about the relation between the treatment and the behavior can be made at this point, however, if changes in the dependent variable are *not* observed in the presence of the independent variable. If accurate measures of the dependent variable were obtained with a measurement system sensitive to changes in the behavior and the treatment condition was implemented as planned, then no behavior change in the presence of the independent variable constitutes disconfirmation of the consequent (B was shown not to be true), and the independent variable may be eliminated as a controlling variable. However, eliminating a treatment from the ranks of controlling variables on the basis of no observed effects presupposes experimental control of the highest order (Johnston & Pennypacker, 1993a).

However, the data set in Figure 7.3 shows that behavior changed in the presence of the independent variable, revealing a correlation between the independent variable and the behavior change. To what extent was the behavior change a function of the independent variable? To pursue this question, the behavior analyst employs the next component of baseline logic: verification.

Verification

The experimenter can increase the probability that an observed change in behavior was functionally related to the introduction of the independent variable by verifying the original prediction of unchanging baseline measures. **Verification** can be accomplished by demonstrating that the prior level of baseline responding would have remained unchanged had the independent variable not been introduced. If that can be demonstrated, this operation verifies the accuracy of the original prediction of continued stable baseline responding and reduces the probability that some uncontrolled (confounding) variable was responsible for the observed change in behavior. Again, the reasoning behind affirmation of the consequent is the logic that underlies the experimental strategy.

Figure 7.4 illustrates the verification of effect in our hypothetical experiment. When steady state responding was established in the presence of the independent variable, the investigator removed the treatment variable, thereby returning to the previous baseline conditions. This tactic allows the possibility of affirming two different antecedent–consequent statements. The first statement and its affirmation follow this pattern:

1. If the independent variable is a controlling factor for the behavior (A), then its removal will coincide with changes in the response measure (B).
2. Removal of the independent variable is accompanied by changes in the behavior (B is true).
3. Therefore, the independent variable controls responding (therefore, A is true).

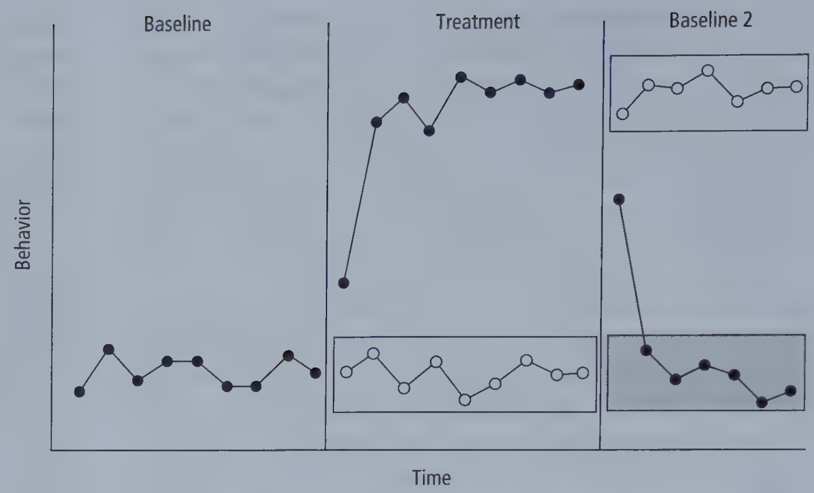


Figure 7.4 Verification of a previously predicted level of baseline responding by withdrawal of the treatment variable. The measures obtained during Baseline 2 (solid data points within the shaded box) show a successful verification and a second affirmation of the consequent based on a comparison with the predicted level of responding (open dots in Baseline 2) in the continued presence of the treatment variable.

The second statement and affirmation follows this pattern:

1. If the original baseline condition controlled the behavior (A), then a return to baseline conditions will result in similar levels of responding (B).
2. The baseline condition is reinstated, and levels of responding similar to those obtained during the original baseline phase are observed (B is true).
3. Therefore, the baseline condition controlled the behavior both then and now (therefore, A is true).

The six measures within the shaded area obtained during Baseline 2 of our hypothetical experiment verify the prediction made for Baseline 1. The open data points in Baseline 2 represent the predicted level of responding if the independent variable had not been removed. (The prediction component of baseline logic applies to steady state responding obtained during any phase of an experiment, baseline and treatment conditions alike.) The difference between the data actually obtained during Treatment (solid data points) and the data obtained during Baseline 2 (solid data points) affirms the first if-A-then-B statement: If the treatment is a controlling variable, then its removal will result in changes in behavior. The similarity between measures obtained during Baseline 2 and those obtained during Baseline 1 confirms the second if-A-then-B statement: If baseline conditions controlled the behavior before, reinstating baseline conditions will result in similar levels of responding.

As before, the observed changes in behavior associated with the introduction and withdrawal of the independent variable are subject to interpretations other than the existence of a functional relation between the two events. However, the case for a functional relation is becoming stronger. When the independent variable was applied, behavior change was observed; when the independent variable was withdrawn, behavior again changed and responding returned to baseline levels. To the extent that the experimenter has effectively controlled the presence and absence of the independent variable and held constant all other variables in the experimental setting that may have influenced the behavior, a functional relation appears likely. An important behavior change has been produced and reversed by the introduction and withdrawal of the independent variable. The process of verification reduces the likelihood that a variable other than the independent variable was responsible for the observed behavior changes.

Does this two-step strategy of prediction and verification constitute sufficient demonstration of a functional relation? What if some uncontrolled variable covaried with the independent variable as it was presented and withdrawn and that uncontrolled variable was responsible for the observed changes in behavior? If such were the case, claiming a functional relation between the target behavior and the independent variable would at best be wrong and at the worst perhaps end the search for the actual controlling variables whose identification and control would contribute to an effective and reliable technology of behavior change.

The appropriately skeptical investigator (and practitioner/research consumer) will also question the reliability of the obtained effect. Was the apparent functional relation a fleeting, one-time-only phenomenon, or will repeated applications of the independent variable reliably (i.e., consistently) produce a similar pattern of behavior change? An effective (i.e., convincing) experimental design yields data responsive to these important questions. To investigate uncertain reliability, the behavior analyst employs the final, and perhaps the most important, component of baseline logic and experimental design: replication.

Replication

> Replication is the essence of believability.
>
> —Baer, Wolf, and Risley (1968, p. 95)

Within the context of any given experiment, **replication** means repeating independent variable manipulations conducted previously in the study and obtaining similar outcomes.[11] Replication within an experiment has two important purposes. First, replicating a previously observed behavior change reduces the probability that some factor other than the independent variable was responsible for the now twice-observed behavior change. Second, replication demonstrates the reliability of the

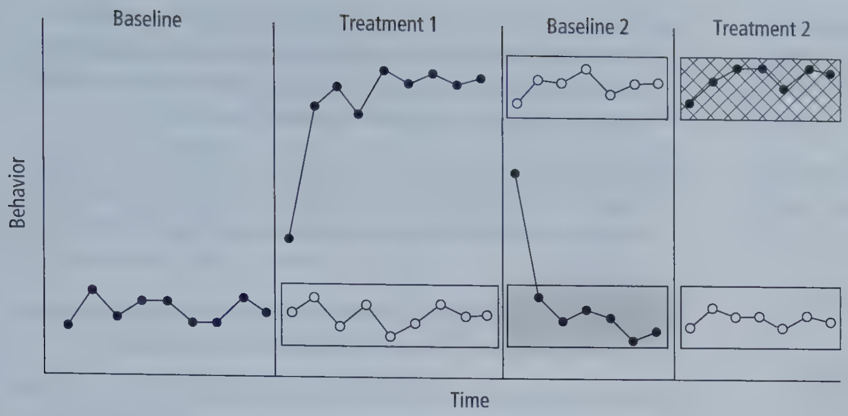


Figure 7.5 Replication of experimental effect accomplished by reintroducing the treatment variable. The measures obtained during Treatment 2 (solid data points within the cross hatched box) enhance the case for a functional relation between the treatment variable and the target behavior.

behavior change; it can be made to happen again. Together, these joint outcomes make replication the basis for inferences that the independent variable is responsible for the observed behavior change.

Figure 7.5 adds the component of replication to our hypothetical experiment. After steady state responding was obtained during Baseline 2, the independent variable is reintroduced. To the extent the data obtained during the Treatment 2 phase (data points within the box shaded with cross-hatched lines) resemble the measures obtained during Treatment 1, replication has occurred. Our hypothetical experiment has now produced powerful evidence of a functional relation between the independent and dependent variables. Confidence in the assertion of a functional relation rests on numerous factors, some of the most important of which are the accuracy and sensitivity of the measurement system; the degree of control the experimenter maintained over all relevant variables; the duration of experimental phases; the stability of responding within each phase; and the speed, magnitude, and consistency of behavior change between conditions. If each of these considerations is satisfied, then replication of effect becomes perhaps the most critical factor in claiming a functional relation.

An independent variable can be manipulated in an effort to replicate an effect many times within an experiment. The number of replications required to demonstrate a functional relation convincingly is related to many considerations, including all of those just enumerated, and to the existence of other similar experiments that have produced the same effects.

SUMMARY

Introduction

1. Measurement can show whether, when, and how much behavior changes, but measurement alone cannot reveal why the behavior changes occurred.
2. A useful technology of behavior change depends upon knowing that specific arrangements of environmental variables will reliably produce desired forms of behavior change.
3. An experimental analysis must be performed to determine how a given behavior functions in relation to specific environmental events.

Concepts and Assumptions Underlying the Analysis of Behavior

4. The overall goal of science is to achieve an understanding of the phenomena under study—socially important behavior, in the case of applied behavior analysis.
5. Science produces understanding at three levels: description, prediction, and control.
6. Descriptive research yields a collection of facts about the observed events—facts that can be quantified and classified.
7. A correlation exists when two events systematically covary with one another. Predictions can be made about the probability that one event will occur based on the occurrence of the other event.

Experimental Control: The Path to and Goal of Behavior Analysis

8. The greatest potential benefits of science are derived from the third, and highest, level of scientific understanding, which comes from establishing experimental control.
9. Experimental control is achieved when a predictable change in behavior (the dependent variable) is reliably produced by the systematic manipulation of some aspect of the environment (the independent variable).
10. An analysis of a behavior has been achieved when a reliable functional relation between the behavior and some specified aspect of the environment has been demonstrated convincingly.
11. A functional analysis does not eliminate the possibility that the behavior under investigation is also a function of other variables.

12. An experiment that shows convincingly that changes in behavior are a function of the independent variable and not the result of uncontrolled or unknown variables has internal validity.

13. External validity refers to the degree to which a study's results are generalizable to other subjects, settings, and/or behaviors.

Behavior: Defining Features and Assumptions that Guide Its Analysis

14. Confounding variables exert unknown or uncontrolled influences on the dependent variable.

15. Because behavior is an individual phenomenon, the experimental strategy of behavior analysis features within-subject (or single-case) methods of analysis.

16. Because behavior is a dynamic and continuous phenomenon that occurs in and changes through time, measuring behavior as it changes over time is a fundamental feature of applied behavior analysis.

17. The assumption of determinism guides experimental method in behavior analysis.

18. Experimental methods in behavior analysis also derive from the assumption that behavioral variability is extrinsic to the organism, not an inherent trait; that is, variability is imposed by environmental variables.

19. Behavior analysts strive to isolate and experimentally manipulate the environmental factors responsible for variability.

Components of Applied Behavior Analysis Experiments

20. A research question specifies what the investigator wants to learn from conducting the experiment; all aspects of an experiment's design and conduct should follow from it.

21. Experiments in applied behavior analysis are referred to as *single-case* research designs because the experimental logic for detecting and analyzing behavior change employs each subject as his or her own control.

22. The dependent variable in an applied behavior analysis experiment is a measurable dimensional quantity of the target behavior.

23. Some studies include multiple dependent variables for three major reasons: (a) To provide additional data paths that serve as controls for evaluating and replicating the effects of an independent variable as it is applied sequentially to each behavior, (b) to assess the presence and extent of the independent variable's effects on behaviors other than the response class to which it was applied directly, and (c) to determine whether changes in the behavior of a person other than the subject occur during the course of an experiment and whether such changes might in turn explain observed changes in the subject's behavior.

24. In addition to precise manipulation of the independent variable, the behavior analyst must hold constant all other aspects of the experimental setting—extraneous variables—to prevent unplanned environmental variation.

25. When unplanned events or variations occur in the experimental setting, the behavior analyst must either wait out their effects or incorporate them into the design of the experiment.

26. Observation and measurement procedures must be conducted in a standardized manner throughout an experiment.

27. Because behavior is a continuous and dynamic phenomenon, ongoing visual inspection of the data is necessary to identify changes in level, trend, and/or variability as they develop during the course of an experiment.

28. Changes in the independent variable are made to maximize its effect on the target behavior.

29. The term *experimental design* refers to the sequence and form of independent variable manipulations in a study.

30. Although an infinite number of experimental designs are possible as a result of the many ways independent variables can be manipulated and combined, an experimenter can make only two basic kinds of independent variable changes: introduce a new condition, or reinstate an old condition.

31. A parametric study compares the differential effects of different values of the independent variable.

32. The fundamental rule of experimentation: change only one variable at a time.

33. Rather than follow rigid, pro forma experimental designs, behavior analysts should select experimental tactics suited to the original research questions, while standing ready to "explore relevant variables by manipulating them in an improvised and rapidly changing design" (Skinner, 1966, p. 21).

Steady State Strategy and Baseline Logic

34. Steady state responding enables a form of inductive reasoning called baseline logic. Baseline logic entails three elements: prediction, verification, and replication.

35. The most common method for evaluating the effects of a given variable is to impose it on an ongoing measure of behavior obtained in its absence. These preintervention data serve as the baseline against which any observed changes in behavior when the independent variable is introduced can be compared.

36. A baseline condition does not necessarily mean the absence of instruction or treatment per se, only the absence of a specific independent variable of experimental interest.

37. In addition to the primary purpose of establishing a baseline as an objective basis for evaluating the effects of the independent variable, three other reasons for baseline data collection are: (a) Systematic observation of the target

behavior prior to intervention sometimes yields information about antecedent-behavior-consequent correlations that may be useful in planning an effective intervention; (b) baseline data can provide valuable guidance in setting initial criteria for reinforcement; and (c) baseline data may reveal that the behavior targeted for change does not warrant intervention.

38. Four types of baseline data patterns are stable, ascending, descending, and variable.
39. The independent variable should be introduced when stable baseline responding has been achieved.
40. The independent variable should not be introduced when an ascending or descending baseline indicates improving performance.
41. The independent variable should be introduced when an ascending or a descending baseline indicates deteriorating performance.
42. An independent variable should not be imposed on a highly variable, unstable baseline.

Affirmation of the Consequent

43. Prediction of future behavior can be made when repeated measures of behavior under relatively constant environmental conditions show little or no variation.
44. Practice effects refer to improvements in performance resulting from opportunities to emit the behavior that must be provided to obtain repeated measures.
45. Extended baseline measurement is not necessary for behaviors that have no logical opportunity to occur.
46. The predictive power of steady state responding enables a form of inductive reasoning known as affirmation of the consequent.
47. Although the logic of affirmation of the consequent is flawed (some other event may have caused the change in behavior), an effective experimental design confirms several if-A-then-B possibilities, each one reducing the likelihood that factors other than the independent variable are responsible for the observed changes in behavior.
48. Verification is accomplished by demonstrating that the prior level of baseline responding would have remained unchanged had the independent variable not been introduced.
49. Replication within an experiment means reproducing a previously observed behavior change by reintroducing the independent variable. Replication demonstrates the reliability of the behavior change and reduces the probability that some factor other than the independent variable was responsible for the now twice-observed behavior change.

KEY TERMS

A–B design
affirmation of the consequent
ascending baseline
baseline
baseline logic
confounding variable
dependent variable
descending baseline
experimental control
experimental design
external validity
extraneous variable
independent variable
internal validity
parametric analysis
practice effects
prediction
research question
replication
single-case designs
stable baseline
steady state responding
steady state strategy
variable baseline
verification

MULTIPLE-CHOICE QUESTIONS

1. Which of the following is not an assumption underlying the analysis of behavior?
   a. Determinism
   b. Empiricism
   c. Philosophic doubt
   d. Pseudoscience
   Hint: (See "Concepts and Assumptions Underlying the Analysis of Behavior")

2. The assumption that the universe is a lawful and orderly place and that natural phenomena occur in relation to other events describes which assumption underlying the analysis of behavior?
   a. Empiricism
   b. Determinism
   c. Parsimony
   d. Philosophic doubt
   Hint: (See "Concepts and Assumptions Underlying the Analysis of Behavior")

3. The three levels of scientific understanding include:
   a. Empiricism, determinism, and philosophic doubt
   b. Verification, prediction, and replication
   c. Group designs, single-subject designs, and correlations
   d. Description, prediction, control

   Hint: (See "Concepts and Assumptions Underlying the Analysis of Behavior")

4. ____ science yields a collection of facts about the observed events—facts that can be quantified and classified.
   a. Replicated
   b. Experimental
   c. Predictive
   d. Descriptive

   Hint: (See "Concepts and Assumptions Underlying the Analysis of Behavior")

5. The highest level of scientific understanding results from establishing:
   a. Experimental control
   b. Prediction
   c. Replication
   d. Analysis

   Hint: (See "Experimental Control: The Path to and Goal of Behavior Analysis")

6. An ________ of behavior has been achieved when a reliable functional relation between the behavior and the independent variable has been demonstrated convincingly.
   a. Application
   b. Investigation
   c. Analysis
   d. Experimental design

   Hint: (See "Experimental Control: The Path to and Goal of Behavior Analysis")

7. ________ refers to the degree to which a study's results are generalizable to other subjects, settings, and/or behaviors.
   a. Internal validity
   b. External validity
   c. Experimental control
   d. Prediction

   Hint: (See "Experimental Control: The Path to and Goal of Behavior Analysis")

8. An experiment has a high degree of _______ when the experiment shows convincingly that changes in behavior are a function of the independent variable and not the result of uncontrolled or unknown variables.
   a. Internal validity
   b. External validity
   c. Replication
   d. Prediction

   Hint: (See "Experimental Control: The Path to and Goal of Behavior Analysis")

9. ________ are variables known or suspected to exert an uncontrolled influence on the dependent variable.
   a. Independent variables
   b. Investigated variables
   c. Confounding variables
   d. Specified variables

   Hint: (See "Experimental Control: The Path to and Goal of Behavior Analysis")

10. Behavior analysts believe behavioral variability is ________ to the organism.
    a. Internal
    b. External
    c. Inevitable
    d. Confined

    Hint: (See "Behavior: Defining Features and Assumptions that Guide Its Analysis")

11. _________ is the "hallmark of applied behavior analysis."
    a. Internal and external validity
    b. Replication, prediction, and control
    c. Single-measurement
    d. Systematic repeated measurement of behavior

    Hint: (See "Behavior: Defining Features and Assumptions that Guide Its Analysis")

12. _________ is the interaction between an organism and its environment.
    a. Measurement
    b. Independent variable
    c. Behavior
    d. Science

    Hint: (See "Behavior: Defining Features and Assumptions that Guide Its Analysis")

13. Which of the following is not considered to be an essential component of experiments in applied behavior analysis?
    a. At least one subject
    b. At least one setting
    c. At least one large group for statistical comparisons
    d. At least one independent variable
    Hint: (See "Components of Applied Behavior Analysis Experiments")

14. A well-planned experiment begins with
    a. A research question
    b. Selection of participants
    c. Selection of a setting
    d. Selection of an experimental design
    Hint: (See "Components of Applied Behavior Analysis Experiments")

15. As a general rule, the independent variable should be introduced when ________ baseline responding has been achieved.
    a. Variable
    b. Replicable
    c. Stable
    d. Complete
    Hint: (See "Steady State Strategy and Baseline Logic")

16. When the data show no evidence of an upward or downward trend and all the measures fall within a small range of values, the data pattern is referred to as
    a. Ascending
    b. Descending
    c. Level
    d. Stable
    Hint: (See "Steady State Strategy and Baseline Logic")

17. The three elements of baseline logic include the following:
    a. Prediction, verification, and replication
    b. Ascending, descending, and stable
    c. Internal validity, external validity, and reliability
    d. Dependent variable, independent variable, baseline
    Hint: (See "Steady State Strategy and Baseline Logic")

ESSAY-TYPE QUESTIONS

1. Discuss the basic assumption scientists hold about the universe.
   Hint: (See "Behavior: Defining Features and Assumptions That Guide Its Analysis")
2. Discuss how experimental control, the highest level of scientific understanding, is achieved in a single-subject research design.
   Hint: (See "Experimental Control: The Path to and Goal of Behavior Analysis")
3. Discuss how behavior analysts approach variability in their data.
   Hint: (See "Behavior: Defining Features and Assumptions That Guide Its Analysis")
4. List the essential components of experiments in behavior analysis.
   Hint: (See "Components of Applied Behavior Analysis Experiments")
5. What provides the basis for comparison in single-subject experimental designs?
   Hint: (See "Components of Applied Behavior Analysis Experiments")
6. Discuss two reasons an experimenter might measure multiple behaviors within one single-subject study.
   Hint: (See "Components of Applied Behavior Analysis Experiments")
7. Discuss the importance of the following "fundamental rule of experimentation" as it relates to the demonstration of a functional relationship between a dependent and an independent variable: "Change only one variable at a time."
   Hint: (See "Intervention or Treatment: Independent Variable")
8. Discuss one scientific benefit and one applied benefit of establishing a baseline level of responding.
   Hint: (See "Steady State Strategy and Baseline Logic")

9. Draw the following from the written description provided:

a. Baseline: High level, stable, no trend

b. Baseline: Low level, stable, no trend

c. Baseline: Slightly increasing, high variability

d. Baseline: Decreasing variable trend

Hint: (See "Steady State Strategy and Baseline Logic")

10. List and describe the three components of baseline logic (also known as "experimental reasoning") used in single-subject research designs.

Hint: (See "Steady State Strategy and Baseline Logic")

NOTES

1. Behavior analysis has benefited immensely from two particularly noteworthy contributions to the literature on experimental method: Sidman's *Tactics of Scientific Research* (1960) and Johnston and Pennypacker's *Strategies and Tactics of Behavioral Research* (1980). Both books are essential reading and working references for any serious student or practitioner of behavior analysis.
2. The researcher's audience (journal editors, reviewers, readers, and practitioners) ultimately determines whether a claimed functional relation is believable or convincing. We examine the believability of research findings further in Chapter 10.
3. Some investigators also use between-group designs to produce results they believe will have more external validity. Group comparison and within-subject experimental methods are compared in Chapter 10.
4. It has become commonplace to refer to the person(s) whose behavior is the dependent variable in an experiment as a *participant*, instead of the more traditional term, *subject*. We use both terms in this text and urge readers to consider Sidman's (2002) perspective on the issue: "[W]e are no longer permitted to call our subjects 'subjects.' The term is supposed to be dehumanizing, and so we are supposed to call them 'participants.' I think this is completely misguided. Experimenters, too, are participants in their experiments. What does making them non-participants do to our perception of science and of scientists? Are experimenters merely robots who follow prescribed and unbreakable scientific rules? Are they supposed just to manipulate variables and coldly record the results of their manipulations? Separating them as nonparticipating manipulators and recorders of the behavior of participants really dehumanizes not only experimenters but, along with them, the whole scientific process." (p. 9)
5. A study by Rindfuss, Al-Attrash, Morrison, and Heward (1998) provides a good example of the extent to which the term *single-subject* research can be a misnomer. A reversal design was used to evaluate the effects of response cards on the quiz and exam scores of 85 students in 5 eighth-grade American history classes. Although a large group of students participated, the study actually consisted of 85 individual experiments; or 1 experiment and 84 replications!
6. Iverson (2013) points out two good reasons why $N = 1$ design should be avoided as a synonym for single-case design: (a) $N = 1$ clearly means the use of only one subject; and (b) in statistics, N denotes the number of measures or data points collected, not the number of subjects. Additionally, *single-case designs* should not be confused with *case studies*, which do involve just one participant and may or may not entail an experimental design.
7. This is the distinguishing feature of the *multiple baseline across behaviors design*, an experimental tactic used widely in applied behavior analysis. Multiple baseline designs are described in Chapter 9.
8. This analytic tactic is known as a *multiple baseline across settings design*. Multiple baseline designs are presented in Chapter 9.
9. How many different experimental designs are there? Because an experiment's design includes careful selection and consideration of each component discussed here (i.e., subject, setting, behavior, etc.), not counting the direct replication of experiments, one could say that there are as many experimental designs as there are experiments.
10. The applied researcher must guard very carefully against assuming automatically that unwanted variability is a function of variables beyond his capability or resources to isolate and control, and thus fail to pursue the investigation of potentially important functional relations.
11. Replication also refers to the repeating of experiments to determine the reliability of functional relations found in previous studies and the extent to which those findings can be extended to other subjects, settings, and/or behaviors (i.e., generality or external validity). The replication of experiments is examined in Chapter 10.

CHAPTER 8

Reversal and Multielement Designs

LEARNING OBJECTIVES

- Analyze the effects of a variety of independent variables using two types of experimental tactics widely used by applied behavior analysts: the reversal design and the alternating treatments design.
- Select the appropriate experimental tactic based on the research question of interest and the appropriateness of the design.
- Discuss how the reversal design and the alternating treatments design and their variations incorporate the elements of baseline logic (prediction, verification, and replication).
- State and describe advantages and disadvantages in using the reversal design and alternating treatments design.
- Identify practical and ethical considerations in using the reversal design and alternating treatments design.

This chapter describes the reversal and multielement designs, two single-case research tactics used widely by applied behavior analysts. In a reversal design, the effects of introducing, withdrawing (or "reversing" the focus of), and reintroducing an independent variable are observed on the target behavior. In a multielement design, two or more experimental treatments are rapidly alternated (or presented simultaneously and the participant is able to select the treatment), and the differential effects on the target behavior noted. We explain how each design incorporates the three elements of steady state strategy—prediction, verification, and replication—and present examples illustrating the major variations of each. Considerations for selecting and using reversal and multielement designs are also presented.

REVERSAL DESIGN

A **reversal design** entails repeated measures of behavior in a given setting that requires at least three consecutive phases: (a) an initial baseline phase in which the independent variable is absent, (b) an intervention phase during which the independent variable is introduced and remains in contact with the behavior, and (c) a return to baseline conditions accomplished by withdrawal of the independent variable. In the widely used notation system for describing experimental designs in applied behavior analysis, the capital letters *A* and *B* denote the first and second conditions, respectively, in a study. Baseline (A) data are collected until steady state responding is achieved. Next, an intervention (B) condition is applied that signifies the presence of a treatment—the independent variable. An experiment entailing one reversal is described as an **A-B-A design**. Although studies using an A-B-A design are reported in the literature (e.g., Armendariz & Umbriet, 1999; Christle & Schuster, 2003; Geller, Paterson, & Talbott, 1982; Jacobson, Bushell, & Risley, 1969; Raiff & Dallery, 2010), the **A-B-A-B design** is preferred because reintroducing the B condition enables replication of treatment effects, which strengthens the demonstration of experimental control (see Figure 8.1).

The A-B-A-B reversal is the most straightforward and generally most powerful within-subject design for demonstrating a functional relation between an environmental manipulation and a behavior. When a reversal design[1] reveals a functional relation, the data show how the behavior works.

> As explanations go, the one offered by the reversal design was not at all a bad one. In answer to the question, "How does this response work?" we could point out demonstrably that it worked like so [e.g., see Figure 8.1]. Of course, it might also work in other ways; but, we would wait until we had seen the appropriate graphs before agreeing to any other way. (Baer, 1975, p. 19, words in brackets added)

Baer's point must not be overlooked: Showing that a behavior "works" in a predictable and reliable way in the presence and absence of a given variable provides only one answer to the question, How does this behavior work? There may be (and quite likely are) other controlling variables for the targeted response class. Whether additional experimentation exploring those other possibilities is necessary or desirable depends on the social and scientific importance of obtaining a more complete analysis.

Operation and Logic of the Reversal Design

Risley (2005) described the rationale and operation of the reversal design as follows:

> The reversal or ABAB design that Wolf reinvented from Claude Bernard's early examples in experimental medicine entailed establishing a baseline of repeated quantified observations sufficient to see a trend and forecast that trend into the near future (A); to then alter conditions and see

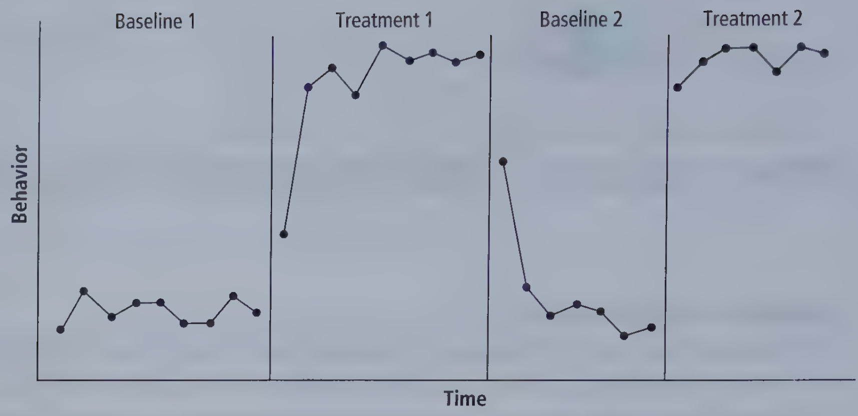


Figure 8.1 Graphic prototype of an A-B-A-B reversal design (hypothetical data).

> if the repeated observations become different than they were forecast to be (B); to then change back and see if the repeated observations return to confirm the original forecast (A); and, finally, to reintroduce the altered conditions and see if the repeated observations again become different than forecast (B). (pp. 280–281)[2]

Because the reversal design was used in Chapter 7 to illustrate baseline logic, a brief review of the roles of prediction, verification, and replication in the reversal design will suffice here. Figure 8.2 shows the same data from Figure 8.1 with the addition of open data points representing predicted measures of behavior if conditions in the previous phase had remained unchanged. After a stable pattern of responding, or counter-therapeutic trend, is obtained in Baseline 1, the independent variable is introduced. In our hypothetical experiment, the measures obtained during Treatment 1, when compared with those from Baseline 1 and with the measures *predicted* by Baseline 1, show that behavior change occurred and that the change in behavior coincided with the intervention. After steady state responding is attained in Treatment 1, the independent variable is withdrawn and baseline conditions re-established. If the level of responding in Baseline 2 parallels or closely approximates the measures obtained during Baseline 1, *verification* of the prediction made from Baseline 1 data is obtained. Stated otherwise, had the intervention not been introduced and had the initial baseline conditions remained in effect, the predicted data path would have appeared as shown in Baseline 2. When withdrawal of the independent variable results in a reversal of the behavior change associated with its introduction, a case builds that the intervention is responsible for the observed behavior change. If reintroduction of the independent variable in Treatment 2 reproduces the behavior change observed during Treatment 1, *replication* of effect has been achieved, and a functional relation demonstrated. Stated in other terms, had the intervention continued and had the second baseline condition not been introduced, the predicted data path of the treatment would have appeared as shown in Treatment 2.

Lambert, Cartledge, Lo, and Heward (2006) used an A-B-A-B design to evaluate the effects of response cards on disruptive behavior and academic responding of students during math lessons in two fourth-grade classrooms. During the baseline (A) condition, single-student responding (SSR), the teacher called on one student who had raised a hand to answer the question. During the treatment (B) condition, response cards (RC), each student could respond to every teacher-posed question by writing an answer on a white laminated board. Figure 8.3 shows the results for disruptive behavior by four students targeted in one of the two classrooms. The data reveal a clear functional relation between the use of response cards and substantial reductions in disruptive behavior. The A-B-A-B design revealed an equally clear functional relation between response cards and increased academic responding.

In the 1960s and early 1970s, applied behavior analysts relied almost exclusively on the A-B-A-B reversal design. The straightforward A-B-A-B design played such a dominant role in the early years of applied behavior analysis that it came to symbolize the field (Baer, 1975). This was no doubt due, at least in part, to the reversal design's potential to expose variables

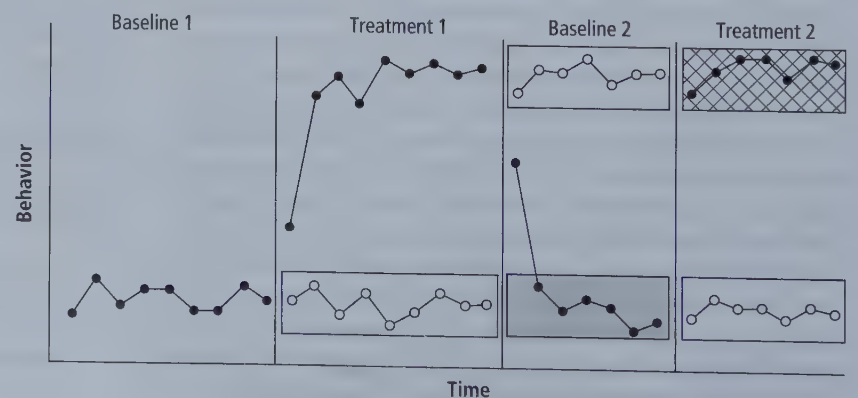


Figure 8.2 Illustration of baseline logic in an A-B-A-B reversal design (hypothetical data). *Prediction:* Open data points (in clear boxes) represent predicted measures if conditions from previous phases remained in effect. *Verification:* Baseline 2 data (lightly shaded box) verify the prediction from Baseline 1. *Replication:* Treatment 2 data (cross-hatch shaded box) replicate the experimental effect.

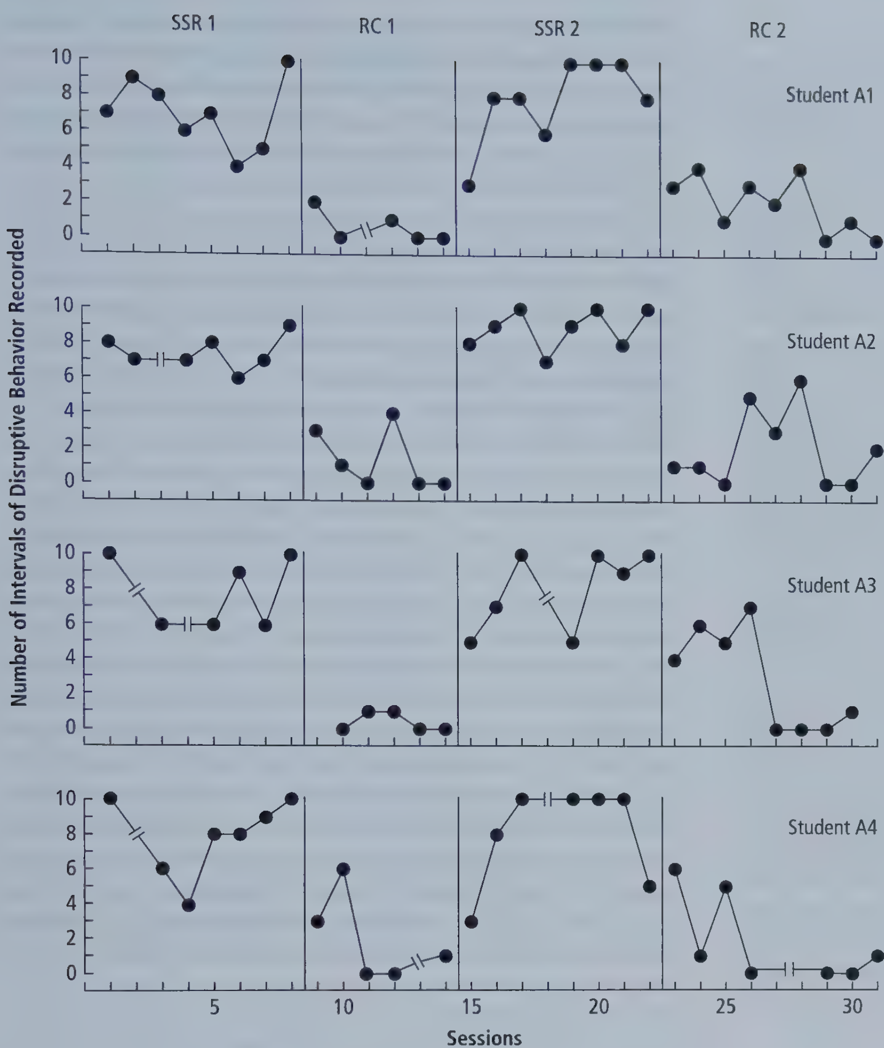


Figure 1. Number of disruptive behaviors during single-student responding (SSR) and response card (RC) conditions for the target students in Classroom A. Breaks in data points represent student absences.

Figure 8.3 Example of an A-B-A-B reversal design.

Based on "Effects of Response Cards on Disruptive Behavior and Academic Responding by Fourth-Grade Urban Students" by M. C. Lambert, G. Cartledge, Y. Lo, and W. L. Heward, 2006, *Journal of Positive Behavioral Interventions*, *8*, p. 93. Copyright 2006 by Sage Publishing.

for what they are—strong and reliable or weak and unstable. Another reason for the reversal design's early dominance may have been that few alternative analytic tactics were available that effectively combined the intrasubject experimental elements of prediction, verification, and replication.

Although the reversal design is just one of many experimental designs available to applied behavior analysts today, the simple, unadorned A-B-A-B design continues to play a major role in the behavior analysis literature (e.g., Beavers, Iwata, & Gregory, 2014; Bicard, Ervin, Bicard, & Baylot-Casey 2012; Hine, Ardoin, & Foster, 2015; Kamps, Conklin, & Wills, 2015; Krentz, Miltenberger, & Valbuena, 2016). Additional examples of the reversal design can be viewed in studies described in this text: Anderson and Long (2002, Figure 24.2); Austin and Bevan (2011, Figure 25.10); Cammilleri and Hanley (2005, Figure 30.14); Cowdery, Iwata, and Pace (1990, Figure 25.6); Deaver, Miltenberger, and Stricker (2001, Figure 24.3); Hall et al. (1971, Figure 14.1); Kuhn, Lerman, Vorndran, and Addison (2006, Figure 23.13); and Lindberg, Iwata, Kahng, and DeLeon (1999, Figure 25.9).

Variations of the A-B-A-B Design

Many applied behavior analysis studies use variations or extensions of the A-B-A-B design.

Repeated Reversals

Perhaps the most obvious variation of the A-B-A-B reversal design is a simple extension in which the independent variable is withdrawn and reintroduced once again (A-B-A-B-A-B). Each additional presentation and withdrawal that reproduces the previously observed effects on behavior increases the likelihood that the behavior changes result from manipulating the independent variable. All other things being equal, an experiment that incorporates multiple reversals presents a more convincing and compelling demonstration of a functional relation than does an experiment with one reversal (e.g., Krentz et al., 2016 [Figure 6.2]; Steege et al., 1990). That said, it is also possible to reach a point of redundancy beyond which the findings of a given analysis are no longer enhanced significantly by additional reversals.

B-A-B and B-A-B-A Designs

The **B-A-B design** begins with the application of the independent variable: the treatment. After stable responding has been achieved during the initial treatment phase (B), the independent variable is withdrawn. If the behavior worsens in the absence of the independent variable (the A condition), the treatment variable is reintroduced in an attempt to recapture the level of responding obtained during the first treatment phase, which would verify the prediction based on the data path obtained during the initial treatment phase.

Practitioners may prefer the B-A-B design to the A-B-A design in that the study ends with the treatment variable in effect. However, in terms of demonstrating a functional relation, the B-A-B design is the weaker of the two because it does not enable an assessment of the effects of the independent variable on the preintervention level of responding. The nonintervention (A) condition in a B-A-B design cannot verify a prediction of a previous nonexistent baseline. Practitioners can remedy this weakness by withdrawing and then reintroducing the independent variable, as in a B-A-B-A-B design (e.g., Dixon, Benedict, & Larson, 2001 [see Figure 25.1]).

Because the B-A-B design provides no data to determine whether the measures of behavior taken during the A condition represent preintervention performance, sequence effects cannot be ruled out: The level of behavior observed during the A condition may have been influenced by the fact that the treatment condition preceded it. Nevertheless, there are exigent situations in which initial baseline data cannot be collected. For instance, the B-A-B design may be appropriate with target behaviors that result in physical harm or danger to the participant or to others. In such instances, withholding a possibly effective treatment until a stable pattern of baseline responding can be obtained may present an ethical problem. For example, Murphey, Ruprecht, Baggio, and Nunes (1979) used a B-A-B design to evaluate the effectiveness of mild punishment combined with reinforcement on the number of self-choking responses by a 24-year-old man with profound intellectual disabilities. After treatment was in effect for 24 sessions, it was withdrawn for 3 sessions, during which an immediate and large increase in self-choking was recorded (see Figure 8.4). Reintroduction of the treatment package reproduced behavior levels observed during the first treatment phase. The mean number of self-chokes during each phase of the B-A-B study was 22, 265, and 24, respectively.

Despite the impressive reduction of behavior, the results of Murphey and colleagues' (1979) B-A-B design study may have been enhanced by gathering and reporting objectively measured data on the level of behavior prior to the first intervention. Presumably, they chose not to collect an initial baseline for ethical and practical reasons. Murphey and colleagues reported anecdotally that self-chokes averaged 434 per day immediately prior to their intervention when school staff had used a different procedure to reduce the self-injurious behavior. This anecdotal information increased the believability of the functional relation suggested by the experimental data from the B-A-B design.

Other situations in which a B-A-B or B-A-B-A design might be warranted include (a) when a treatment is already in place (e.g., Marholin, Touchette, & Stuart, 1979; Pace & Troyer, 2000), (b) when "baseline levels of performance are obvious because the behavior has never been performed (e.g., exercise for many of us, practicing a musical instrument, reading, eating

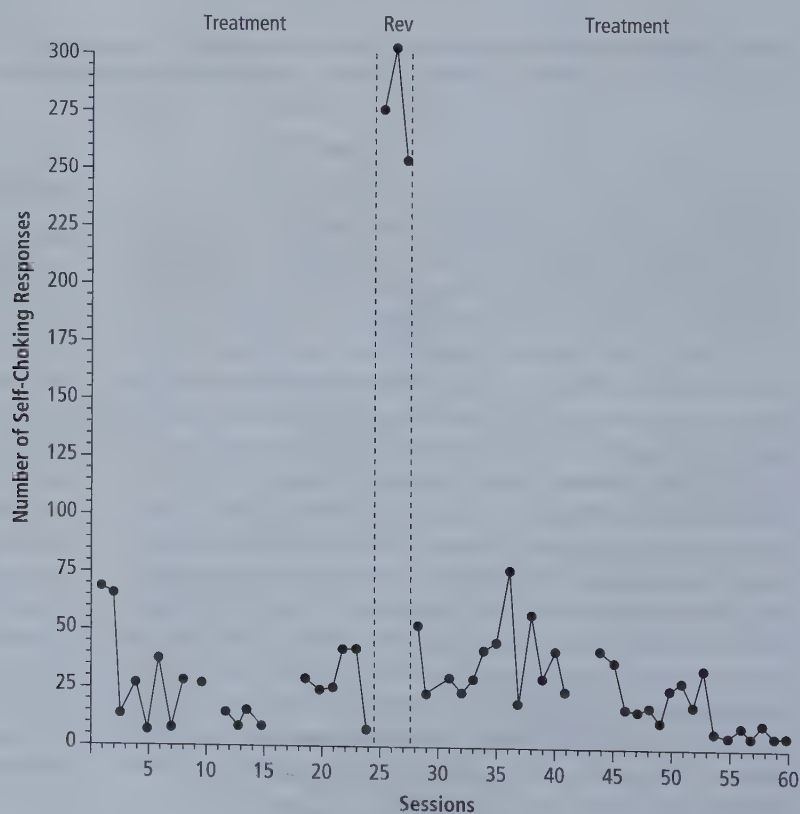


Figure 8.4 Example of a B-A-B reversal design.

Based on "The Use of Mild Punishment in Combination with Reinforcement of Alternate Behaviors to Reduce the Self-Injurious Behavior of a Profoundly Retarded Individual" by R. J. Murphy, M. J. Ruprecht, P. Baggio, and D. L. Nunes, 1979, *AAESPH Review, 4*, p. 191. Copyright 1979 by the *AAESPH Review*.

healthful foods)" (Kazdin, 2011, p. 134), (c) when the behavior analyst has limited time in which to demonstrate practical and socially significant results (e.g., Robinson et al., 1981), and (d) when the client or significant others want an intervention to begin immediately (e.g., Kazdin & Polster, 1973).

Multiple Treatment Reversal Designs

Experiments that use the reversal design to compare the effects of two or more experimental conditions to baseline and/or to one another are said to use a **multiple treatment reversal design**. The letters *C*, *D*, and so on, denote when additional conditions were in effect during the study, as in the following designs: A-B-A-C-B-C (Dickman, Bright, Montgomery, & Miguel, 2012); A-B-C-A-C (Kamps et al., 2015); A-B-C-A-C-B-C (Falcomata, Roane, Hovanetz, Kettering, & Keeney (2004); A-B-A-B-C-A-C-B (Hansen & Wills, 2014); and A-B-A-C-A-D-A-E-A-F (Wilson & Tunnard, 2014).

These designs are considered variations of the reversal design because they embody the experimental method and logic of the reversal tactic: Responding in each phase provides baseline (or control condition) data for the subsequent phase (prediction), independent variables are withdrawn in an attempt to reproduce levels of behavior observed in a previous condition (verification), and each independent variable that contributes fully to the analysis is introduced at least twice (replication). Independent variables can be introduced, withdrawn, changed in value, combined, and otherwise manipulated to produce an endless variety of experimental designs.

For example, Kennedy and Souza (1995) used an A-B-C-B-C-A-C-A-C design to analyze and compare the effects of two kinds of competing sources of stimulation on eye poking by a 19-year-old student with profound disabilities. Geoff had a 12-year history of poking his forefinger into his eyes during periods of inactivity, such as after lunch or while waiting for the bus. The two treatment conditions were music (B) and a video game (C). During the music condition, Geoff was given a Sony Walkman radio with headphones. The radio was tuned to a station that his teacher and family thought he preferred. Geoff had continuous access to the music during this condition, and he could remove the headphones at any time. During the video game condition, Geoff was given a small handheld video game on which he could observe a variety of visual patterns and images on the screen with no sound. As with the music condition, Geoff had continuous access to the video game and could discontinue using it at any time.

Figure 8.5 shows the results of the study. Following an initial baseline phase (A) in which Geoff averaged 4 eye pokes per hour, the music condition (B) was introduced and eye pokes decreased to a mean of 2.8 per hour. The video game (C) was implemented next, and eye pokes decreased further to 1.1 per hour. Measures obtained during the next two phases—a reintroduction of music (B) followed by a second phase of the video game (C)—replicated previous levels of responding under each condition. This B-C-B-C portion of the experiment revealed a functional relation between video game condition and lower frequency of eye pokes compared to music. The final five phases of the experiment (C-A-C-A-C) provided an experimental comparison of the video game and baseline (no-treatment) condition.

In most instances, extended designs involving multiple independent variables are not preplanned. Instead of following a predetermined structure that dictates when and how experimental manipulations will be made, the behavior analyst makes design decisions based on ongoing assessment of the data. For example, Cummings and Carr (2005) used an A-B-C-A-D-BC-A-E-F-A-BF-BCDEF design to evaluate the effects of five interventions that resulted in a treatment package (the BCDEF condition) successful in eliminating the joint dislocation behavior of a child with autism.

> [A] single experiment may be viewed as a number of successive designs that are collectively necessary to clarify relations between independent and dependent variables. Thus, some design decisions might be made in response to the data unfolding as the investigation progresses. This sense of design encourages the experimenter to pursue in more dynamic fashion the solutions to problems of experimental control immediately upon their emergence. (Johnston & Pennypacker, 1980, pp. 250–251)

Students of applied behavior analysis should not interpret this description of experimental design as a recommendation for a free-form approach to the manipulation of independent variables. The researcher must always obey the rule of changing only one variable at a time and must understand the opportunities for

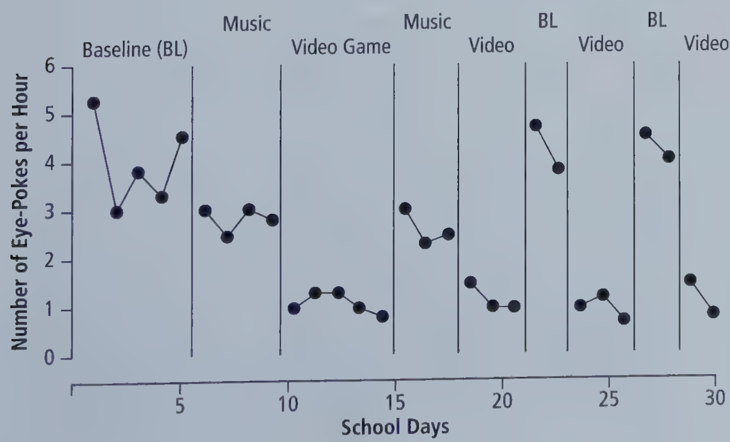


Figure 8.5 Example of a multiple treatment reversal design (A-B-C-B-C-A-C-A-C).

From "Functional Analysis and Treatment of Eye Poking," by C. H. Kennedy and G. Souza, 1995, *Journal of Applied Behavior Analysis*, *28*, p. 33. Copyright 1995 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

legitimate comparisons and the limitations that a given sequence of manipulations places on the conclusions that can be drawn from the results.

Experiments using a reversal design to compare two or more treatments are vulnerable to confounding by sequence effects. **Sequence effects** are the effects on a subject's behavior in a given condition that are influenced by the subject's experience with a prior condition. For example, the analyst must use caution in interpreting results from the A-B-C-B-C design that is produced by the following fairly common sequence of events in practice: After baseline (A), an initial treatment (B) is implemented and little or no behavioral improvements are noted. A second treatment (C) is next tried, and the behavior improves. A reversal is then conducted by reintroducing the first treatment (B), followed by reinstatement of the second treatment (C) (e.g., Foxx & Shapiro, 1978). In this case, we can only speak knowingly about the effects of C when it follows B. Recapturing the original baseline levels of responding before introducing the second treatment condition (i.e., an A-B-A-C-A-C sequence) reduces the threat of sequence effects (or helps to expose them for what they are).

An A-B-A-B-C-B-C design, for instance, enables direct comparisons of B to A and C to B, but not of C to A. An experimental design consisting of A-B-A-B-B+C-B-B+C (e.g., Jason & Liotta, 1982) permits an evaluation of the additive or interactive effects of B+C, but does not reveal the independent contribution of C. And in both of these examples, it is impossible to determine what effects, if any, C may have had on the behavior if it had been implemented prior to B. Manipulating each condition so that it precedes and follows every other condition in the experiment (e.g., A-B-A-B-C-B-C-A-C-A-C) is the only way to know for sure. However, manipulating multiple conditions requires significant time and resources, and such extended designs become more susceptible to confounding by maturation and other historical variables not controlled by the experimenter.

Noncontingent Reinforcement (NCR) Reversal Technique

With interventions based on positive reinforcement, it can be hypothesized that observed improvements in behavior result from the participant's feeling better about himself because of the enriched environment created by the reinforcement, not because specific instances of the target behavior have been immediately followed by contingent reinforcement. This hypothesis is most often advanced when interventions consisting of social reinforcement are involved. For example, a person may claim that it doesn't matter *how* the teacher's praise and attention were given; the student's behavior improved because the praise and attention created a warm and supporting environment. If, however, the behavioral improvements observed during a contingent reinforcement condition are lost during a condition in which equal amounts of the same consequence are delivered independent of the occurrence of the target behavior, a functional relation between the reinforcement contingency and behavior change may be demonstrated.

This experimental control tactic, called **noncontingent reinforcement (NCR) reversal technique**, can show that behavior change is the result of *contingent* reinforcement, not simply the presentation of or contact with the stimulus event (Thompson & Iwata, 2005). A study by Baer and Wolf (1970a) on the effects of teachers' social reinforcement on the cooperative play of a preschool child provides an excellent example of the NCR reversal technique (Figure 8.6). The authors described the use and purpose of the design as follows:

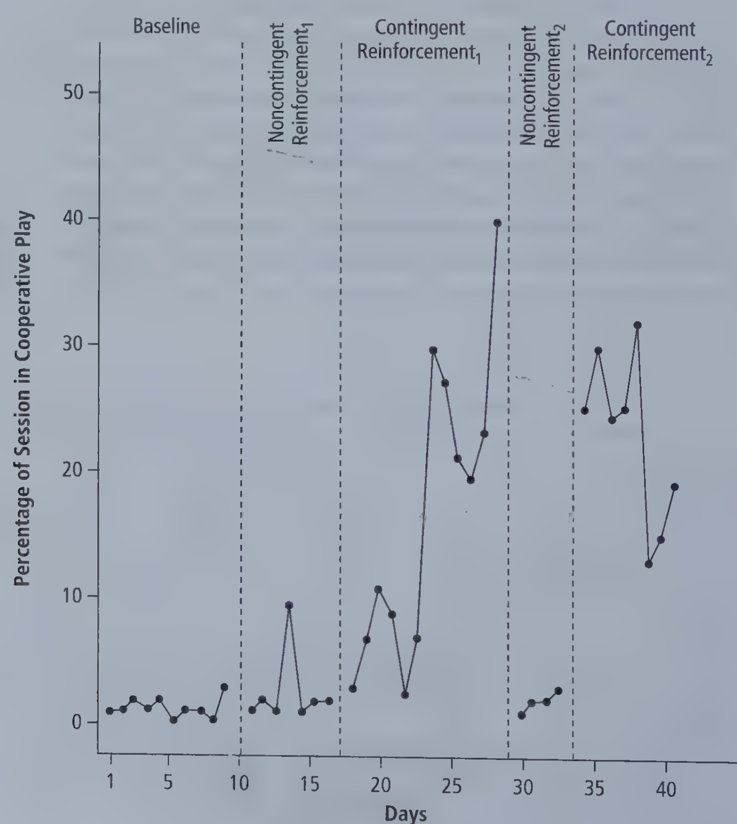


Figure 8.6 Reversal design using noncontingent reinforcement (NCR) as a control technique.

From "Recent Examples of Behavior Modification in Pre-school Settings," by D. M. Baer and M. M. Wolf, in *Behavior Modification in Clinical Psychology*, pp. 14–15, edited by C. Neuringer and J. L. Michael, 1970, Upper Saddle River, NJ: Prentice Hall. Copyright 1970 by Prentice Hall. Adapted by permission.

> [The teachers first collected] baselines of cooperative and other related behaviors of the child, and of their own interaction with the child. Ten days of observation indicated that the child spent about 50% of each day in proximity with other children (meaning within 3 feet of them indoors, or 6 feet outdoors). Despite this frequent proximity, however, the child spent only about 2% of her day in cooperative play with these children. The teachers, it was found, interacted with this girl about 20% of the day, not all of it pleasant. The teachers, therefore, set up a period of intense social reinforcement, offered not for cooperative play but free of any response requirement at all: the teachers took turns standing near the girl, attending closely to her activities, offering her materials, and smiling and laughing with her in a happy and admiring manner. The results of 7 days of this noncontingent extravagance of social reinforcement were straightforward: the child's cooperative play changed not at all, despite the fact that the other children of the group were greatly attracted to the scene, offering the child nearly double the chance to interact with them cooperatively. These 7 days having produced no useful change, the teachers then began their planned reinforcement of cooperative behavior. . . . Contingent social reinforcement, used in amounts less than half that given during the noncontingent period, increased the child's cooperative play from its usual 2% to a high of 40% in the course of 12 days of reinforcement. At that point, in the interests of certainty, the teachers discontinued contingent reinforcement in favor of noncontingent. In the course of 4 days, they lost virtually all of the cooperative behavior they had gained during the reinforcement period of the study, the child showing about a 5% average of cooperative play over that period of time. Naturally, the study concluded with a return to the contingent use of social reinforcement, a recovery of desirable levels of cooperative play, and a gradual reduction of the teacher's role in maintaining that behavior. (pp. 14–15)

Using NCR as a control condition to demonstrate a functional relation is advantageous when it is not possible or appropriate to eliminate completely the event or activity used as a contingent reinforcement. For example, Lattal (1969) employed NCR as a control condition to "reverse" the effects of swimming as reinforcement for tooth brushing by children in a summer camp. In the contingent reinforcement condition, the campers could go swimming only if they had brushed their teeth; in the NCR condition, swimming was available whether or not tooth brushing occurred. The campers brushed their teeth more often in the contingent reinforcement condition.

The usual procedure is to deliver NCR on a fixed or variable time schedule independent of the subject's behavior. A potential weakness of the NCR control procedure becomes apparent when high rate of the desired behavior has been produced during the preceding contingent reinforcement phase. It is probable in such situations that at least some instances of NCR, delivered according to a predetermined time schedule, will follow occurrences of the target behavior closely in time, and thereby function as adventitious, or accidental, reinforcement (Thompson & Iwata, 2005). In fact, an intermittent schedule of reinforcement might be created inadvertently that results in even higher levels of performance than those obtained under contingent reinforcement. (Intermittent schedules of reinforcement and their effects are described in Chapter 13). In such cases the investigator might consider using one of the two control techniques described next, both of which involve "reversing" the behavioral focus of the contingency.[3]

DRO Reversal Technique

One way to ensure that reinforcement will not immediately follow the target behavior is to deliver reinforcement immediately following the subject's performance of *any behavior other than the target behavior.* With a **DRO reversal technique**, the control condition consists of delivering the event suspected of functioning as reinforcement following the emission of any behavior other than the target behavior (e.g., Baer, Peterson, & Sherman, 1967 Osborne, 1969; Poulson, 1983). For example, Reynolds and Risley (1968) used contingent teacher attention to increase the frequency of talking in a 4-year-old girl enrolled in a preschool program for disadvantaged children. After a period of teacher attention contingent on verbalization, in which the girl's talking increased from a baseline average of 11% of the intervals observed to 75%, a DRO condition was implemented during which the teachers attended to the girl for any behavior except talking. During the 6 days of DRO, the girl's verbalization dropped to 6%. Teacher attention was then delivered contingent on talking, and the girl's verbalization "immediately increased to an average of 51%" (p. 259).

DRI/DRA Reversal Technique

During the control condition in a **DRI/DRA reversal technique**, occurrences of a specified behavior that is either incompatible with the target behavior (i.e., the two behaviors cannot possibly be emitted at the same time) or an alternative to the target behavior are immediately followed by the same consequence previously delivered as contingent reinforcement for the target behavior. Goetz and Baer's (1973) investigation of the effects of teacher praise on preschool children's creative play with building blocks illustrates the use of a DRI control condition. Figure 8.7 shows the number of different block forms (e.g., arch, tower, roof, ramp) constructed by the three children who participated in the study. During baseline (data points indicated by the letter *N*), "the teacher sat by the child as she built with the blocks, watching closely but quietly, displaying neither criticism nor enthusiasm about any particular use of the blocks" (p. 212). During the next phase (the *D* data points), "the teacher remarked with interest, enthusiasm, and delight every time that the child placed and/or rearranged the blocks so as to create a form that had not appeared previously in that session's construction(s). . . . 'Oh, that's very nice—that's different!'" (p. 212). Then, after increasing form diversity was clearly established, instead of merely withdrawing verbal praise and returning to the initial baseline condition, the teacher provided descriptive praise only when the children had constructed the same forms (the *S* data points). "Thus, for the next two to four sessions, the teacher continued to display interest, enthusiasm, and delight, but only at those times when the

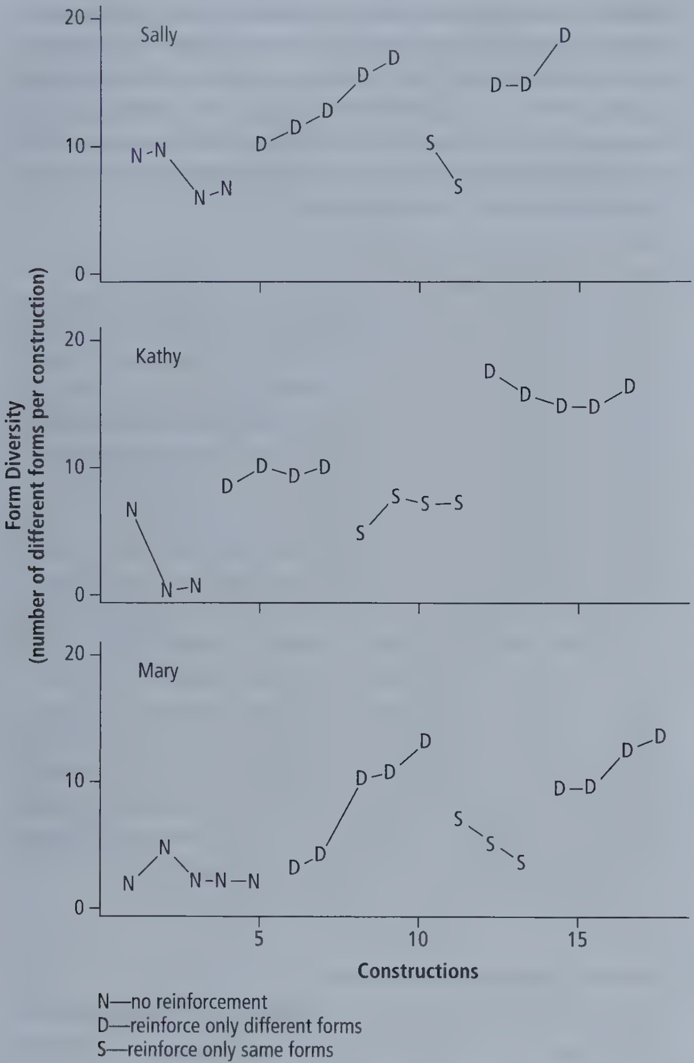


Figure 8.7 Reversal design using a DRI control technique.

Based on "Social Control of Form Diversity and the Emergence of New Forms in Children's Blockbuilding" by E. M. Goetz and D. M. Baer, 1973, *Journal of Applied Behavior Analysis*, 6, p. 213. Copyright 1973 by the Society for the Experimental Analysis of Behavior, Inc.

child placed and/or rearranged a block so as to create a repetition of a form already apparent in that session's construction(s). . . . Thus, no first usage of a form in a session was reinforced, but every second usage of that form and every usage thereafter within the session was. . . . 'How nice—another arch!'" (p. 212). The final phase of the experiment entailed a return to descriptive praise for different forms. Results show that the form diversity of children's block building was a function of teacher praise and comments. The DRI reversal tactic allowed Goetz and Baer to determine that it was not just the delivery of teacher praise and comment that resulted in more creative block building by the children; the praise and attention had to be contingent on different forms to produce increasing form diversity.[4]

Considering the Appropriateness of the Reversal Design

The primary advantage of the reversal design is its ability to provide a clear demonstration of the existence (or absence) of a functional relation between the independent and dependent variables. An investigator who reliably turns the target behavior on and off by presenting and withdrawing a specific variable makes a clear and convincing demonstration of experimental control. In addition, the reversal design enables quantification of the amount of behavior change over the preintervention level of responding. And the return to baseline provides information on the need to program for maintenance. Furthermore, a complete A-B-A-B design ends with the treatment condition in place.[5]

In spite of its strengths as a tool for analysis, the reversal design entails some potential scientific and social disadvantages that should be considered prior to its use. The considerations are of two types: irreversibility, which affects the scientific utility of the design; and the social, educational, and ethical concerns related to withdrawing a seemingly effective intervention.

Irreversibility: A Scientific Consideration

A reversal design is not appropriate in evaluating the effects of a treatment variable that, by its very nature, cannot be withdrawn once it has been presented. Although independent variables involving reinforcement and punishment contingencies can be manipulated with some certainty—the experimenter either presents or withholds the contingency—an independent variable such as providing information or modeling, once presented, cannot simply be removed. For example, a reversal design would not be an effective element of an experiment investigating the effects of attending an in-service training workshop for teachers during which participants observed a master teacher use contingent praise and attention with students. After the participants have listened to the rationale for using contingent praise and attention and observed the master teacher model it, the exposure provided by that experience could not be withdrawn. Such interventions are said to be irreversible.

Irreversibility of the dependent variable must also be considered in determining whether a reversal would be an effective analytic tactic. Behavioral **irreversibility** means that a level of behavior observed in an earlier phase cannot be reproduced even though the experimental conditions are the same as they were during the earlier phase (Sidman, 1960). Once improved, many target behaviors of interest to the applied behavior analyst remain at their newly enhanced level even when the intervention responsible for the behavior change is removed. From a clinical or educational standpoint, such a state of affairs is desirable: The behavior change is shown to be durable, persisting even in the absence of continued treatment. For example, Kuoch and Mirenda (2003) reported irreversibility as a positive outcome of their study evaluating the effects of social stories on problem behavior. Problem behavior by three children with autism decreased from initial baseline levels during a social stories intervention and remained low when the intervention was withdrawn, a pattern of results the authors described as "short-term maintenance of behavioral reduction once the social stories were discontinued" (p. 224). However, irreversibility is a problem if demonstration of the independent variable's role in the behavior change depends on verification by recapturing baseline levels of responding.

For example, baseline observations might reveal very low, almost nonexistent, rates of talking and social interaction for a young child. An intervention consisting of teacher-delivered

social reinforcement for talking and interacting could be implemented, and after some time the girl might talk to and interact with her peers at a frequency and in a manner similar to that of her classmates. The independent variable, teacher-delivered reinforcement, could be terminated in an effort to recapture baseline rates of talking and interacting. But the girl might continue to talk to and interact with her classmates even though the intervention, which may have been responsible for the initial change in her behavior, is withdrawn. In this case a source of reinforcement uncontrolled by the experimenter—the girl's classmates talking to and playing with her as a consequence of her increased talking and interacting with them—could maintain high rates of behavior after the teacher-delivered reinforcement is no longer provided. In such instances of irreversibility, an A-B-A-B design would fail to reveal a functional relation between the independent variable and the target behavior.

Nonetheless, one of the major objectives of applied behavior analysis is establishing socially important behavior through experimental treatments so that the behavior will contact natural "communities of reinforcement" to maintain behavioral improvements in the absence of treatment (Baer & Wolf, 1970b). When irreversibility is suspected or apparent, in addition to considering DRO or DRI/DRA conditions as control techniques, investigators can consider other experimental tactics, most notably the multiple baseline designs described in Chapter 9. A multiple treatment reversal design requires a reversible dependent variable. If the initial treatment condition in a multiple treatment reversal design changes an irreversible behavior, the opportunity to assess and compare the effects of any other treatment conditions is lost.

Withdrawing an Effective Intervention: A Social, Educational, and Ethical Consideration

Although it can yield an unambiguous demonstration of experimental control, withdrawing a seemingly effective intervention to evaluate its role in behavior change presents a legitimate cause for concern. One must question the appropriateness of any procedure that allows (indeed, seeks) an improved behavior to deteriorate to baseline levels of responding. Various concerns have been voiced over this fundamental feature of the reversal design. Although considerable overlap exists among the concerns, they can be classified as having primarily a social, educational, or ethical basis.

Social Concerns. Applied behavior analysis is, by definition, a social enterprise. Behaviors are selected, defined, observed, measured, and modified by and for people. Sometimes the people involved in an applied behavior analysis—administrators, teachers, parents, and participants—object to the withdrawal of an intervention they associate with desirable behavior change. Even though a reversal may provide the most unqualified picture of the behavior–environment relation under study, it may not be the analytic tactic of choice because key participants do not want the intervention to be withdrawn. When a reversal design offers the best experimental approach scientifically and poses no ethical problems, the behavior analyst may choose to explain the operation and purpose of the tactic to those who do not favor it. But it is unwise to attempt a reversal without the full support of the people involved, especially those who will be responsible for withdrawing the intervention. Without their cooperation the procedural integrity of the experiment could easily be compromised. For example, people who are against the withdrawal of treatment might sabotage the return to baseline conditions by implementing the intervention, or at least those parts of it they consider the most important.

Educational and Clinical Issues. Educational or clinical issues concerning the reversal design are often raised in terms of instructional time lost during the reversal phases, as well as the possibility that the behavioral improvements observed during intervention may not be recaptured when treatment is resumed after a return to baseline conditions. Prolonged reversals are neither justifiable nor necessary (Stolz, 1978). If preintervention levels of responding are reached quickly, reversal phases can be quite short in duration. Sometimes only three or four sessions are needed to show that initial baseline rates have been reproduced (e.g., Ashbaugh & Peck, 1998; Cowdery et al., 1990 [Figure 25.6]; Hall et al., 1971, Figure 14.1). Two or three brief reversals can provide an extremely convincing demonstration of experimental control. Concern that the improved levels of behavior will not return when the treatment variable is reintroduced, while understandable, has not been supported by empirical evidence. Hundreds of published studies have shown that behavior acquired under a given set of environmental conditions can be reacquired rapidly during subsequent reapplication of those conditions.

Ethical Concerns. A serious ethical concern must be addressed when the use of a reversal design is considered for evaluating a treatment for self-injurious or dangerous behaviors. With mild self-injurious or aggressive behaviors, short reversal phases consisting of one or two baseline probes can sometimes provide the empirical evidence needed to reveal a functional relation (e.g., Akers, Higbee, Pollard, Pellegrino, & Gerencser, 2016; Kelley, Jarvie, Middlebrook, McNeer, & Drabman, 1984; Luce, Delquadri, & Hall, 1980; Murphey et al., 1979 [Figure 8.4]). For example, in their study evaluating a treatment for elopement (i.e., running away from supervision) by a child with attention-deficit hyperactivity disorder (ADHD), Kodak, Grow, and Northup (2004) returned to baseline conditions for a single session (see Figure 8.8).

Nonetheless, with some behaviors it may be determined that withdrawing an intervention associated with improvement for even a few one-session probes would be unethical. In such cases experimental designs that do not rely on the reversal tactic must be used.

MULTIELEMENT DESIGN

An important and frequently asked question by teachers, therapists, and others who are responsible for changing behavior is this: Which of these treatments will be most effective with this student or client? In many situations, the research literature, the analyst's experience, and/or logical extensions of the principles of behavior point to several possible interventions. Determining which of several possible treatments or combination of treatments will produce the greatest improvement in behavior is a

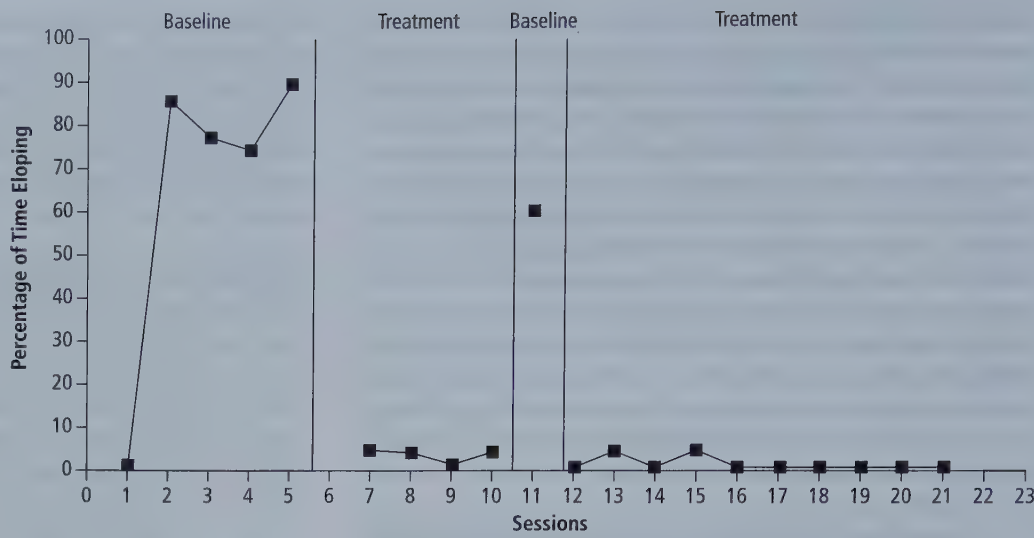


Figure 8.8 Reversal design with a single-session return-to-baseline probe to evaluate and verify effects of treatment for a potentially dangerous behavior.

Based on "Functional Analysis and Treatment of Elopement for a Child with Attention Deficit Hyperactivity Disorder" by T. Kodak, L. Grow, and J. Northrup, 2004, *Journal of Applied Behavior Analysis, 37*, p. 231. Copyright 2004 by the Society for the Experimental Analysis of Behavior, Inc.

primary task for applied behavior analysts. As described earlier, although a multiple treatment reversal design (e.g., A-B-C-B-C) can be used to compare the effects of two or more treatments, such designs have some inherent limitations. Because the different treatments in a multiple treatment reversal design are implemented during separate phases that occur in a particular order, the design is particularly vulnerable to confounding by sequence effects (e.g., Treatment C may have produced its effect only because it followed Treatment B). A second disadvantage of using the reversal tactic to compare multiple treatments is the extended time required to demonstrate differential effects. Most behaviors targeted for change by teachers and therapists are selected because those behaviors need immediate improvement. An experimental design that will quickly reveal the most effective treatment among several possible approaches is important for the applied behavior analyst. The **multielement design** (also called **alternating treatments design**) is an experimentally sound and efficient method for comparing the effects of two or more treatments.

Operation and Logic of the Multielement Design

The multielement design is characterized by the rapid alternation of two or more distinct treatments (i.e., independent variables) while their effects on the target behavior (i.e., dependent variable) are measured. In contrast to the reversal design, in which experimental manipulations are made after steady state responding is achieved in a given phase of an experiment, the different interventions in a multielement design are manipulated independent of the level of responding. To aid the subject's discriminating which treatment condition is in effect at any given time, a distinct stimulus (e.g., a sign, verbal instructions, different-colored worksheets) is often associated with each treatment. Barlow and Hayes (1979) introduced the term *alternating treatments design* and described it this way:

> The data are plotted separately for each intervention to provide a ready visual representation of the effects of each treatment. Because confounding factors such as time of administration have been neutralized (presumably) by counterbalancing, and because the two treatments are readily discriminable by subjects through instructions or other discriminative stimuli, differences in the individual plots of behavior change corresponding with each treatment should be attributable to the treatment itself, allowing a direct comparison between two (or more) treatments. (p. 200)

Figure 8.9 shows a graphic prototype of an alternating treatments design comparing the effects of two treatments, A and B, on some response measure. In an alternating treatments design, the different treatments can be alternated in a variety of ways. For example, the treatments might be (a) alternated across daily sessions, one treatment in effect each day; (b) administered in separate sessions occurring within the same day; or (c) implemented each during a portion of the same session. Counterbalancing the days of the week, times of day, sequence in which the different treatments occur (e.g., first or second each day), persons delivering the different treatments, and so forth, reduces the probability that any observed differences in behavior are the result of variables other than the treatments themselves. For example, assume that Treatments A and B in Figure 8.9 were each administered for a single

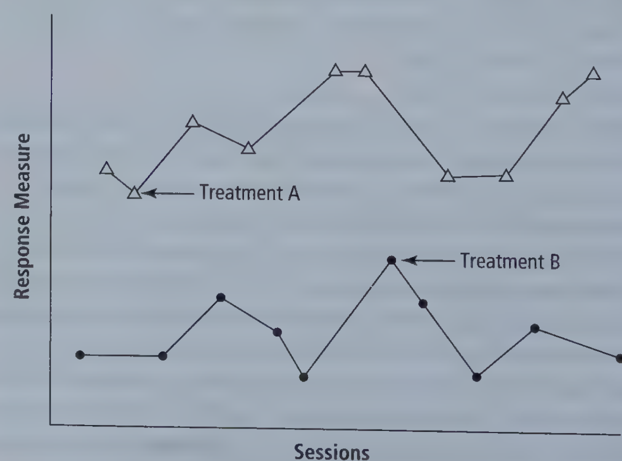


Figure 8.9 Graphic prototype of a multielement design comparing the differential effects of two treatments (A and B).

30-minute session each day, with the daily sequence of the two treatments determined by a coin flip.

The data points in Figure 8.9 are plotted on the horizontal axis to reflect the actual sequence of treatments each day. Thus, the horizontal axis is labeled *Sessions*, and each consecutive pair of sessions occurred on a single day. Some published reports of experiments using an alternating treatments design in which two or more treatments were presented each day (or session) plot the measures obtained during each treatment above the same point on the horizontal axis, thus implying that the treatments were administered simultaneously. This practice masks the temporal order of events and has the unfortunate consequence of making it difficult for the researcher or reader to discover potential sequence effects.

The three components of steady state strategy—prediction, verification, and replication—are found in the alternating treatments design. However, each component is not readily identified with a separate phase of the design. In an alternating treatments design, each successive data point for a specific treatment plays all three roles: It provides (a) a basis for the *prediction* of future levels of responding under that treatment, (b) potential *verification* of the previous prediction of performance under that treatment, and (c) the opportunity for *replication* of previous effects produced by that treatment.

To see this logic unfold, the reader should place a piece of paper over all the data points in Figure 8.9 except those for the first five sessions of each treatment. The visible portions of the data paths provide the basis for predicting future performance under each respective treatment. Moving the paper to the right reveals the two data points for the next day, each of which provides a degree of verification of the previous predictions. As more data are recorded, the predictions of given levels of responding within each treatment are further strengthened by continued verification (if those additional data conform to the same level and/or trend as their predecessors). Replication occurs each time Treatment A is reinstated and measurement reveals responding similar to previous Treatment A measures and different from those obtained when Treatment B is in effect. Likewise, another mini-replication is achieved each time a reintroduction of Treatment B results in measures similar to previous Treatment B measures and different from Treatment A levels of responding. A consistent sequence of verification and replication is evidence of experimental control and strengthens the investigator's confidence of a functional relation between the two treatments and different levels of responding.

The presence and degree of experimental control in a multielement design is determined by visual inspection of the differences between (or among) the data paths representing the different treatments. Inspecting multielement design data should include visual analysis of mean shift in level, trend, and variability throughout the comparison phase(s) of the experiment (Diller, Barry, & Gelino, 2016). Experimental control is defined in this instance as objective, believable evidence that different levels of responding (and perhaps degrees of trends and variability) are predictably and reliably produced by the presence of the different treatments. When the data paths for two treatments show no overlap with each other and either stable levels or opposing trends, a clear demonstration of experimental control has been made. Such is the case in Figure 8.9, in which there is no overlap of data paths and the picture of differential effects is clear. When some overlap of data paths occurs, a degree of experimental control over the target behavior can still be demonstrated if the majority of data points for a given treatment fall outside the range of values of the majority of data points for the contrasting treatment.

The extent of any differential effects produced by two treatments is determined by the vertical distance—or fractionation—between their respective data paths and quantified by the vertical axis scale. The greater the vertical distance, the greater the differential effects of the two treatments on the response measure. It is possible for experimental control to be shown between two treatments but for the amount of behavior change to be socially insignificant. For instance, experimental control may be demonstrated for a treatment that reduces a person's severe self-injurious behavior from 10 occurrences per hour to 2 per hour, but the participant is still engaged in self-mutilation. However, if the vertical axis is scaled meaningfully, the greater the separation of data paths on the vertical axis, the higher the likelihood that the difference represents a socially significant effect.

Data from an experiment that compared the effects of two types of group-contingent rewards on the spelling accuracy of fourth-grade underachievers (Morgan, 1978) illustrate how the alternating treatments design reveals experimental control and the quantification of differential effects. The six children in the study were divided into two equally skilled teams of three on the basis of pretest scores. Each day during the study the students took a five-word spelling test. The students received a list of the words the day before, and a 5-minute study period was provided just prior to the test. Three different conditions were used in the alternating treatments design: (a) *no game*, in which the spelling tests were graded immediately and returned to the students, and the next scheduled activity in the school day began; (b) *game*, in which test papers were graded immediately, and each member of the team who had attained the highest total score received a Certificate of Achievement and was allowed to stand up and cheer; and (c) *game plus*, consisting of the same procedure as the game condition, plus each student on the winning team also received a small trinket (e.g., a sticker or pencil).

The results for Student 3 (see Figure 8.10) show that experimental control over spelling accuracy was obtained between the no-game condition and both the game and the game-plus conditions. Only the first two no-game data points overlap the lower range of scores obtained during the game or the game-plus conditions. However, the data paths for the game and game-plus conditions overlap completely and continuously throughout the study, revealing no difference in spelling accuracy between the two treatments. The vertical distance between the data paths represents the amount of improvement in spelling accuracy between the no-game condition and the game and the game-plus conditions. The mean difference between the two game conditions and the no-game condition was two words per test. Whether such a difference represents a significant improvement is an educational question, not a mathematical or statistical one, but most educators and parents would agree

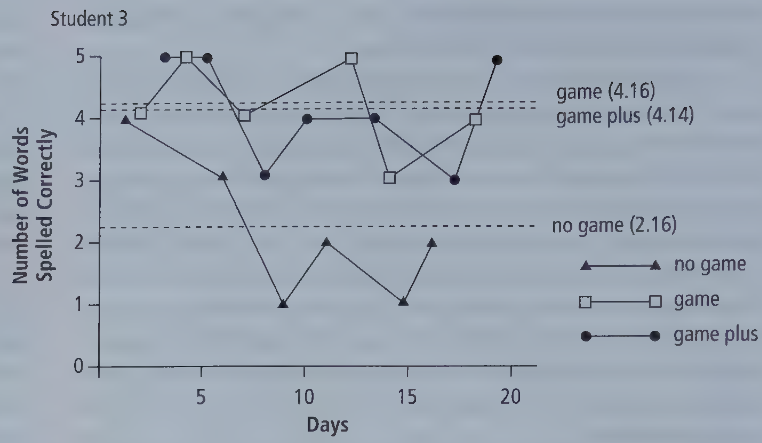


Figure 8.10 A multielement design comparing the effects of three different treatments on the spelling accuracy of a fourth-grade student.

Based on *Comparison of Two "Good Behavior Game" Group Contingencies on the Spelling Accuracy of Fourth-Grade Students* by Q. E. Morgan, 1978, unpublished master's thesis, The Ohio State University.

that an increase of two words spelled correctly out of five is socially significant, especially if that gain can be sustained from week to week. The cumulative effect over a 180-day school year would be impressive. There was virtually no difference in Student 3's spelling performance between the game and game-plus conditions. However, even a larger mean difference would not have contributed to the conclusions of the study because of the lack of experimental control between the game and the game-plus treatments.

Student 6 earned consistently higher spelling scores in the game-plus condition than he did in the game or no-game conditions (see Figure 8.11). Experimental control was demonstrated between the game-plus and the other two treatments for Student 6, but not between the no-game and game conditions. Again, the difference in responding between treatments is quantified by the vertical distance between the data paths. In this case there was a mean difference of 1.55 correctly spelled words per test between the game-plus and no-game conditions.

Figures 8.10 and 8.11 illustrate two other important points about the multielement design. First, the two graphs show how the multielement design enables a quick comparison of interventions. After 20 sessions the teacher had sufficient empirical evidence for selecting the most effective consequences for each student. If only two conditions had been compared, even fewer sessions may have been required to identify the most effective intervention. Second, these data underscore the importance of evaluating treatment effects at the level of the individual subject. All six children spelled more words correctly under one or both of the game conditions than they did under the no-game condition. However, Student 3's spelling accuracy was equally enhanced by either the game or the game-plus contingency, whereas Student 6's spelling scores improved only when a tangible reward was available.

Variations of the Multielement Design

The multielement design can be used to compare one or more treatments to a no-treatment or baseline condition, assess the relative contributions of individual components of a package intervention, and perform parametric investigations in which different values of an independent variable are alternated to determine differential effects on behavior change. The following sections describe the most common variations of the multielement design.

Single-phase Multielement Design

A single-phase experiment in which the effects of two or more treatment conditions are compared is the most basic application of the multielement design (e.g., Barbetta, Heron, & Heward, 1993; McNeish, Heron, & Okyere, 1992; Morton, Heward, & Alber, 1998). For example, Brady and Kubina (2010) compared the multiplication fact fluency of three elementary-age students with attention-deficit hyperactivity disorder (ADHD) under two practice procedures: whole time practice trials and endurance building

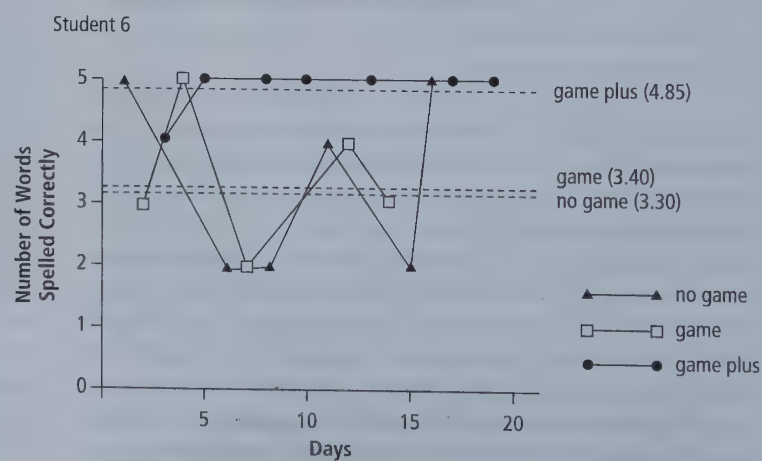


Figure 8.11 A multielement design comparing the effects of three different treatments on the spelling accuracy of a fourth-grade student.

Based on *Comparison of Two "Good Behavior Game" Group Contingencies on the Spelling Accuracy of Fourth-Grade Students* by Q. E. Morgan, 1978, unpublished master's thesis, The Ohio State University.

practice trials. Each procedure lasted 2.5 minutes and consisted of students writing as many answers to single-digit multiplication math fact sheet assessments as they could for a combined total of 2 minutes and receiving corrective feedback from the experimenter for a combined total of 30 seconds. The procedures differed in how time for writing answers and corrective feedback was distributed. The whole time practice condition consisted of one 1-minute practice trial, 30 seconds of corrective feedback, and a 1-minute multiplication math fact sheet assessment. The endurance building practice condition consisted of three 20-second practice trials interspersed with 10-second corrective feedback, and a 1-minute multiplication math fact sheet assessment. The two practice conditions were alternated daily.

Figure 8.12 shows the results on the daily multiplication math fact sheet assessments for one of the three participants. Brady and Kubina (2010) reported that as participants' "performance levels for the three 20 second intervals improved so too did their ability to perform over the longer time interval of 1 minute" (p. 90). Although practice time equaled 1 minute in both conditions, the students practiced 30% more problems during the endurance building practice trials condition than they did in the whole time practice trial condition.

A no-treatment condition is often incorporated into the design as one of the treatments to be compared. For example, the no-game condition in the Morgan (1978) study served as a no-treatment control condition against which the students' spelling scores in the game and game-plus conditions were compared (see Figures 8.10 and 8.11). Including a no-treatment control condition as one of the conditions in a multielement design provides valuable information on any differences in responding under the intervention treatment(s) and no treatment. However, the measures obtained during the no-treatment control condition should not be considered representative of an unknown baseline level of responding. It may be that the measures obtained in the no-treatment condition represent only the level of behavior under a no-treatment condition when it is interspersed within an ongoing series of treatment condition(s), and do not represent the level of behavior that existed before the alternating treatments design was begun.

The single-phase multielement design is the primary experimental tool for conducting functional analyses of problem behavior (Chapter 27). Figure 6.5 and Figures 27.7, 27.8, 25.10, and 27.11 in this text show examples of functional analyses using multielement designs.

Two-phase Design: Baseline and Comparison

Investigators often use a two-phase experimental design in which baseline measures are collected until a stable level of responding or counter-therapeutic trend is obtained prior to the alternating treatments phase (e.g., Kranak, Alber-Morgan, & Sawyer, 2017; Martens, Lochner, & Kelly, 1992 [see Figure 13.6]; Neef, Iwata, & Page, 1980 [see Figure 6.12]). The baseline condition may be continued during the treatments comparison phase as a no-treatment control condition.

Wilder, Atwell, and Wine (2006) used a multielement design with an initial baseline phase to evaluate the relative effectiveness of three levels of treatment integrity on compliance by typically developing preschool children with instructions not usually complied with by the children: "Give me the [snack item]," "Put the toy away," and "Come here." Each session the therapist presented one of the three instructions 10 times at a rate of once per minute. Compliance, defined as doing what the therapist instructed within 10 seconds, resulted in brief praise from the therapist throughout the study. During baseline, noncompliance produced no response from the therapist. During the alternating treatments phase, the therapist responded to noncompliance by implementing a three-step prompting procedure for 100% of trials for one instruction (e.g. "Give me the [snack item]"), 50% of trials for another instruction (e.g., "Put

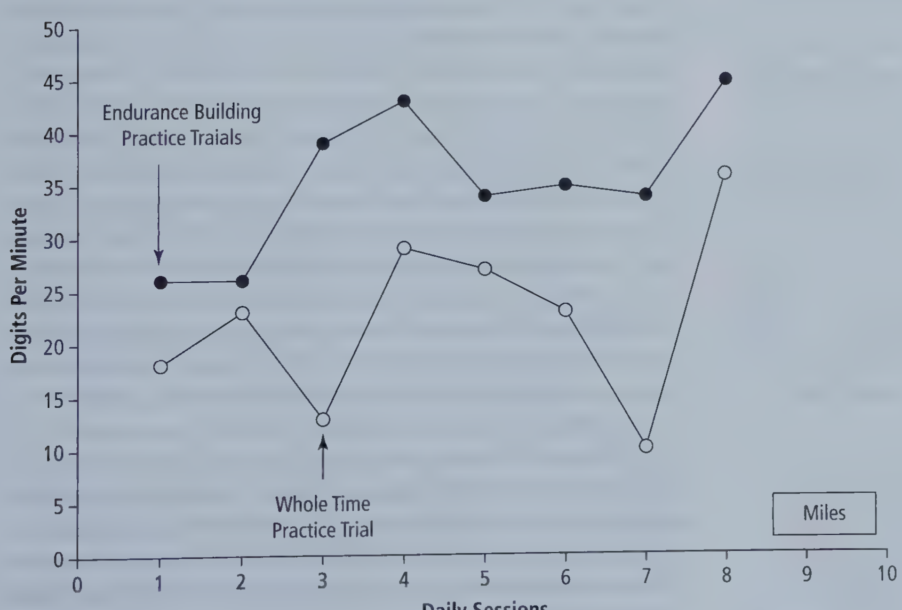


Figure 2. Miles correct per minute frequency during endurance building practice trials and whole time trial conditions

Figure 8.12 A multielement design comparing two treatment conditions.

Based on "Endurance of Multiplication Fact Fluency for Students with Attention Deficit Hyperactivity Disorder" by K. K. Brady and R. M. Kubina, Jr. 2010, *Behavior Modification*, *34*(2), p. 86. Copyright 2010 by Sage Publishing. Reprinted by permission.

the toy away"), and 0% of trials for the third instruction (e.g., "Come here"). The 0% integrity condition was identical to baseline and served as a no-treatment control condition.

Figure 8.13 shows the results for the two children who participated in the study. Each child complied most often with the instruction for which the therapist implemented the prompting intervention with 100% integrity, complied at a higher than baseline level with the instruction associated with 50% treatment integrity, and seldom or never complied with the instruction associated with 0% integrity (i.e., baseline conditions). By associating the 100% integrity condition with a different instruction for each child, the experimenters strengthened the replication of the functional relation between the 100% integrity condition and the highest level of compliance.

Sindelar, Rosenberg, and Wilson (1985) described a variation of the multielement design for comparing the efficiency of instructional procedures they called the **adapted alternating treatments design**. The comparison phase of the design features the alternating application of two (usually) or more different teaching methods, each method applied to different but equivalent sets of instructional items. All items are topographically different members of the same response or skill class, such as reading printed words, defining vocabulary terms, spelling words, answering math problems, and stating history facts. Sindelar and colleagues described the design's operation and purpose as follows:

> A baseline in which the equivalence of performance on the two sets is demonstrated is followed by an experimental condition in which acquisition of one set taught by one method is compared to the acquisition of the second set taught by the second method. Differences are demonstrated when acquisition of one set is more rapid than acquisition of the other, and the effect is consistent across subjects, settings, or behaviors. (p. 70)

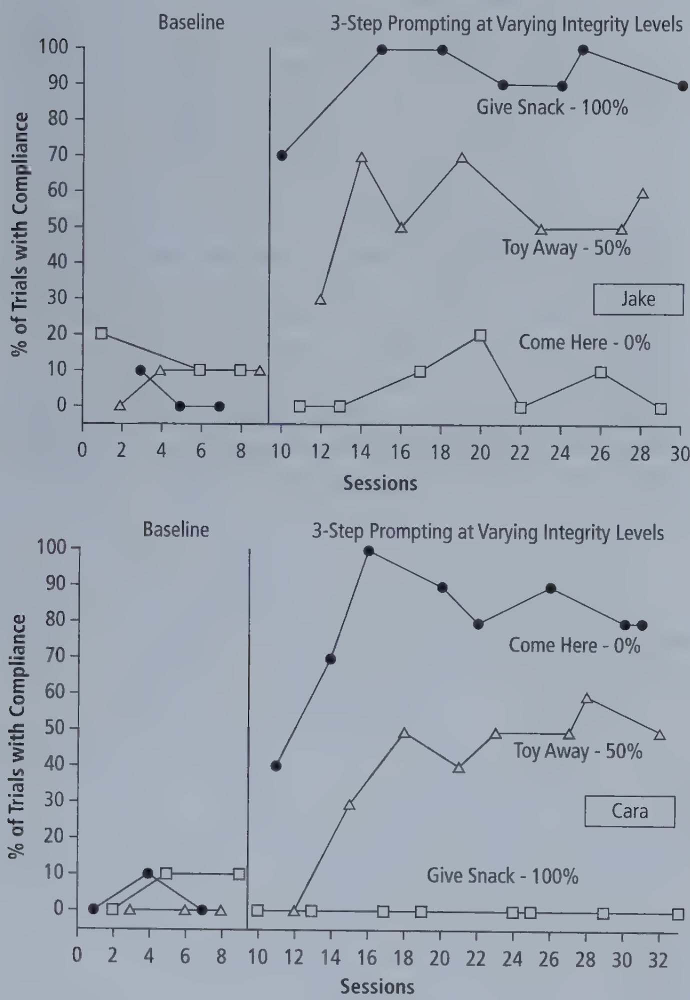


Figure 1. Percentage of trials with compliance across baseline and three-step prompting integrity levels (i.e., 100%, 50%, and 0%).

Figure 8.13 Multielement design with an initial baseline.

Based on "The Effects of Varying Levels of Treatment Integrity on Child Compliance During Treatment with a Three-Step Prompting Procedure" by D. A., Wilder, J. Atwell, and B. Wine, 2006, Journal of *Applied Behavior Analysis*, *39*, p. 372. Copyright 2006 by the Society for the Experimental Analysis of Behavior, Inc.

Vladescu and Kodak (2013) used an adapted alternating treatments design to analyze the extent and speed with which preschool children with autism would acquire secondary learning targets as a function of presenting information on secondary learning targets during the antecedent or consequence portion of learning trials for primary targets. Pretests determined an individualized pool of unknown targets—naming items on picture cards or responding to fill-in-the-blank statements ("The opposite of hot is ____ ?")—for each participant. The targets were assigned to different sets, each set representing an experimental condition (i.e., teaching method) distinguished by whether and when information on secondary targets was presented. The experimenters equated the sets by considering the number of syllables contained in responses to the targets and putting targets that sounded similar in different sets. In the *antecedent condition*, the experimenter presented information on a secondary target (e.g., held up a picture and said, "This is a seal."), waited 3 seconds, and then presented the primary target (e.g., held up a picture of a lion and asked, "What is it?"). In the *consequence condition*, the experimenter presented information on a secondary target in the same way immediately after delivering reinforcement for the participant's response to the primary target. An error correction procedure ensured that primary target trials ended with a correct response by the participant. In the *secondary-targets-only condition*, the experimenter presented information on secondary targets in the absence of teaching primary targets. In the *primary-targets-only condition*, no secondary target information was provided. Participants' acquisition of secondary-target information was measured by probes, conducted every one to three training sessions in other conditions. Participants received no feedback or reinforcement for responses to secondary target presentations in any condition.

Figure 8.14 shows the results for one of the four participants. Rick acquired the primary targets in the antecedent, consequence, and primary-targets-only conditions (top panel) in five, six, and five sessions, respectively. He correctly echoed the secondary targets on 100% of opportunities during the antecedent, consequence, and secondary-targets-only conditions. Probes showed that Rick mastered the secondary targets during the training of primary targets, and he did so in a similar number of

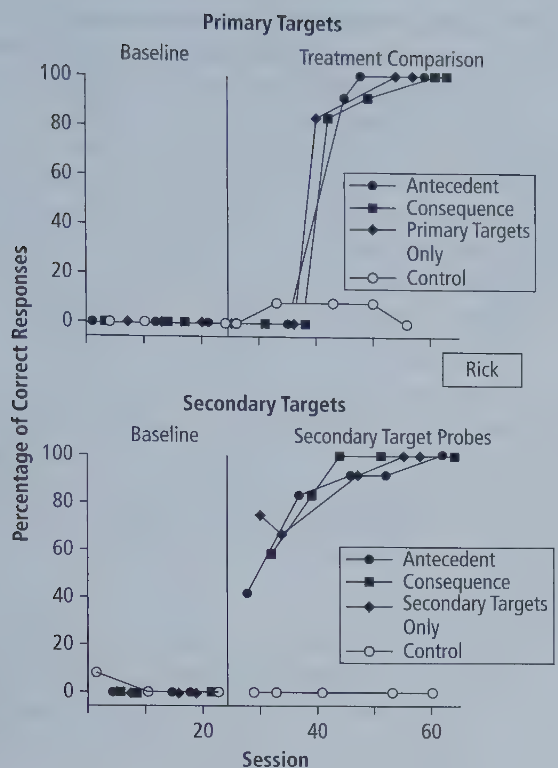


Figure 8.14 Adapted alternating treatments design.

Based on "Increasing Instructional Efficiency by Presenting Additional Stimuli in Learning Trials for Children with Autism Spectrum Disorders" by J. C. Vladescu and T. M. Kodak, 2013, *Journal of Applied Behavior Analysis*, *46*, p. 812. Copyright 2013 by the Society for the Experimental Analysis of Behavior, Inc.

sessions as the primary targets (bottom panel). Like Rick, two of the three other children in the study mastered secondary targets without explicit instruction. That is, when information on secondary targets was presented in the antecedent or consequence condition, the children learned double the number of targets in a similar amount of training time compared to the condition in which only the primary targets were presented.

Three-phase Design: Baseline, Comparison, and Best Treatment

Another widely used variation of the alternating treatments design consists of three sequential phases: an initial baseline phase, a second phase comparing two or more alternating treatments (one of which may be the baseline control condition), and a final phase in which only the most effective treatment is administered (e.g., Heckaman, Alber, Hooper, & Heward, 1998; Mechling, 2006; Ollendick, Matson, Esvelt-Dawson, & Shapiro, 1980).

Tincani (2004) used such a design to investigate the relative effectiveness of sign language and picture exchange training on the acquisition of mands (requests for preferred items) by two children with autism.[6] A related research question was whether a relation existed between students' pre-existing motor imitation skills and their abilities to learn mands through sign language or by picture exchange. Two assessments were conducted for each student prior to baseline. A stimulus preference assessment was conducted to identify a list of 10 to 12 preferred items (e.g., drinks, edibles, toys), and each student's ability to imitate 27 hand, arm, and finger movements similar to those required for sign language was assessed.[7]

The purpose of baseline was to ensure that the participants were not able to request preferred items with picture exchange, sign language, or speech prior to training. Baseline trials consisted of giving the student 10 to 20 seconds of noncontingent access to a preferred item, removing the item briefly, and then placing it out of the student's reach. A laminated 2-inch-by-2-inch picture of the item was placed in front of the student. If the student placed the picture symbol in the experimenter's hand, signed the name of the item, or spoke the name of the item within 10 seconds, the experimenter provided access to the item. If not, the item was removed and the next item on the list was presented. Following a three-session baseline, during which neither participant emitted an independent mand in any modality, the alternating treatments phase was begun.

The sign language training procedures were adapted from Sundberg and Partington's (1998) *Teaching Language to Children with Autism or Other Developmental Disabilities.* The simplest sign from American Sign Language for each item was taught. Procedures used in the Picture Exchange Communication System (PECS) training condition were adapted from Bondy and Frost's (2002) *The Picture Exchange Communication System Training Manual.* In both conditions, training on each preferred item continued for five to seven trials per session, or until the participant showed no interest in the item. At that time, training then began on the next item and continued until all 10 or 12 items on the participant's list of preferred items had been presented. During the study's final phase, each participant received either sign language or PECS training only, depending on which method had been most successful during the alternating treatments phase.

The percentage of independent mands by the two students throughout the study is shown in Figures 8.15 (Jennifer) and 8.16 (Carl). Picture exchange training was clearly more effective than sign language for Jennifer. Jennifer demonstrated weak motor imitation skills in the pre-baseline assessment, correctly imitating 20% of the motor movements attempted in the pre-baseline imitation assessment. After a slight modification in sign language training procedures was implemented to eliminate Carl's prompt dependency, he emitted independent mands more often during sign language training than with picture exchange training. Carl's pre-existing motor imitation skills were better than Jennifer's. He imitated correctly 43% of the attempted motor movements in the pre-baseline imitation assessment.

This study highlights the importance of individual analyses and exploring the possible influence of variables not manipulated during the study. In discussing the study's results, Tincani (2004) noted that

> For learners without hand-motor imitation skills, including many children with autism, PECS training may be more appropriate, at least in terms of initial mand acquisition. Jennifer had weak hand-motor imitation skills prior to intervention and learned picture exchange more rapidly than sign language. For learners who have moderate hand-motor imitation skills, sign language training may be equally, if not more, appropriate. Carl had moderate hand-motor imitation skills prior to intervention and learned sign language more rapidly than picture exchange. (p. 160)

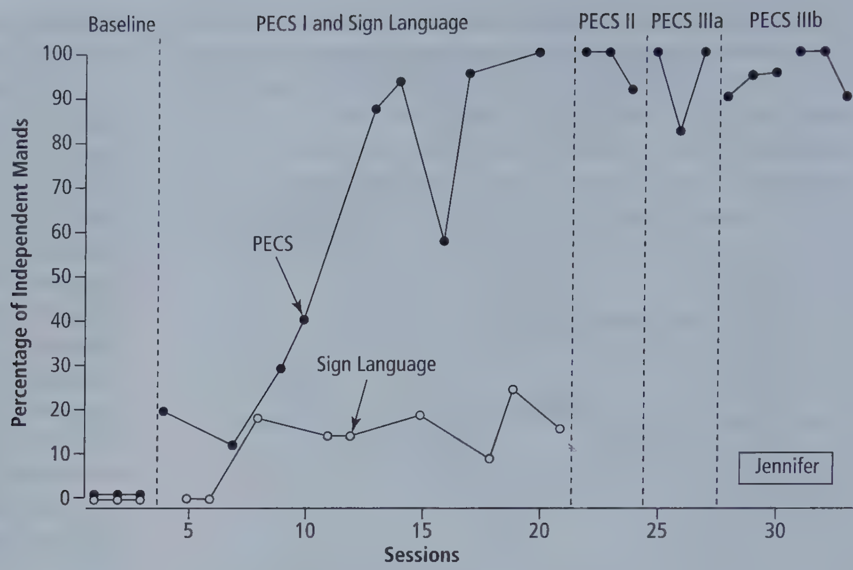


Figure 8.15 Example of a multielement design with initial baseline and a final best-treatment-only condition.

Based on "Comparing the Picture Exchange Communication System and Sign Language Training for Children with Autism" by M. Tincani, 2004, *Focus on Autism and Other Developmental Disabilities, 19,* p. 160. Copyright 2004 by Pro-Ed.

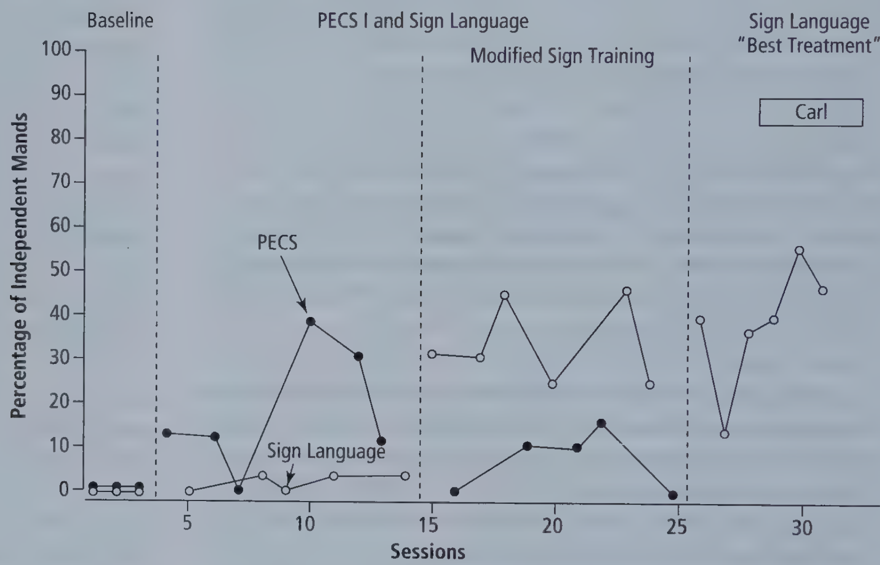


Figure 8.16 Example of a multielement design with initial baseline and a final best-treatment-only condition.

Based on "Comparing the Picture Exchange Communication System and Sign Language Training for Children with Autism" by M. Tincani, 2004, *Focus on Autism and Other Developmental Disabilities, 19,* p. 159. Copyright 2004 by Pro-Ed.

Three-phase Design: Baseline, Comparison, and Participant's Choice

Another three-phase variation of the multielement design entails an initial baseline, a comparison phase in which two or more treatment conditions (one of which may be the baseline control condition) are alternated, and a final "choice" phase in which only the treatment selected by each participant is implemented.

Jowett Hirst, Dozier, and Payne (2016) used this sequence of phases in two experiments assessing (a) the relative effects of the differential reinforcement of alternative behavior (DRA) and response cost (RC) procedures on the on-task behavior of typically developing preschool children and (b) the children's preferences for the procedures.[8] In Experiment I, conducted in the context of small-group activities, both DRA and RC increased on-task behavior for the majority of children and five of six children chose RC more often than DRA. Prior to each session in Experiment II, conducted in the context of individual work tasks (e.g., worksheets with printed letters and shapes for tracing), the experimenter explained the rules and contingencies that would be in effect:

> *Baseline.* "Today you get the white board, and there are no tokens. When we start, you can either work on tracing or play with toys. If you are working (i.e., tracing), nothing will happen. If you are not working, nothing will happen." (p. 338)
>
> *Differential reinforcement of alternative behavior.* "Today you get the green board, and it doesn't have any tokens on it. When we start, you can either work on tracing or play with toys. If you are working, you will get a token; if you are not working, you will not get a token. At the end, you can trade your tokens for prizes and snacks. If you don't have any tokens, you don't get anything." (pp. 338)
>
> *Response cost.* "Today you get the red board, and it has 10 tokens on it. When we start, you can either work on

tracing or play with toys. If you are working, you will keep your tokens; if you are not working, you will lose tokens. At the end,. . . " [same as DRA instructions] (p. 338)

Choice. The experimenter placed the stimuli (i.e., poster and token boards) associated with each type of condition (i.e., baseline, RC, and DRA) near the participant and reminded him or her of the contingencies associated with each set of materials . . . . [T]he experimenter asked the participant to pick (by pointing to or touching a set of materials) which session he or she wanted to do. (pp. 338–339)

On 10 occasions (at intervals ranging from 15 to 45 seconds) during each 5-minute session, except for baseline sessions, the experimenter delivered (DRA) or removed (RC) a token per the condition in effect. Figure 8.17 shows the results for 6 of the 10 children who participated in the multielement comparison of DRA and RC in Experiment II. All children displayed moderate and variable or extremely low levels of on-task behavior in baseline. As in Experiment I, both DRA and RC resulted in roughly equal levels of increased on-task behavior during the comparison phase (Adam was the lone exception), and those levels were maintained during the choice phase. Two participants selected DRA for each session of the choice phase, two selected RC exclusively, and six switched their selections.

Concurrent Chains (or Schedule) Design

A design in which participants choose between two or more concurrently available treatments is termed a **concurrent schedule** or **concurrent chains design**.[9] Participants are presented with two or more response options; each option is associated with a distinctive discriminative stimulus and leads to a different set of treatment procedures. Hanley, Piazza, Fisher, and Maglieri (2005) used a concurrent chains design to assess clients' preference for problem behavior treatments. Participants were two children who had been admitted to an inpatient unit that specialized in treating severe problem behavior. A functional analysis demonstrated that both children engaged in problem behavior (hitting, biting, head butting, throwing objects at people) most often in the attention condition (the therapist provided brief reprimand following each occurrence of problem behavior). The researchers then evaluated (with reversal and multielement designs) the effects of several function-based treatments for each participant. Functional communication training (FCT) was moderately effective in reducing problem behavior for both children, and FCT plus punishment was more effective in reducing problem behavior for both children than FCT alone.

With FCT, each child was trained to emit an alternative response (saying "attention please" or "excuse me" for Betty; handing the therapist a card on which the word "play" was printed for Jay). Emitting the alternative behavior resulted in 20 seconds of attention from the therapist. Occurrences of problem behavior were followed by no differential consequence (i.e., extinction). Procedures for FCT plus punishment were identical to FCT except that each occurrence of problem behavior involved "a 30-s hands-down procedure for Jay (the therapist stood behind Jay and held his hands to his sides) and a 30-s hands-down and visual-screen procedure for Betty (the therapist stood behind the child and placed one arm around the child's arms while placing the other hand over the child's eyes)." (Hanley et al., 2005, p. 56).

The researchers then used a concurrent-chains strategy to evaluate each child's relative preference for various treatments. Prior to entering the therapy room during each session of this analysis, each child was presented with three different-colored microswitches, each switch corresponding to one of

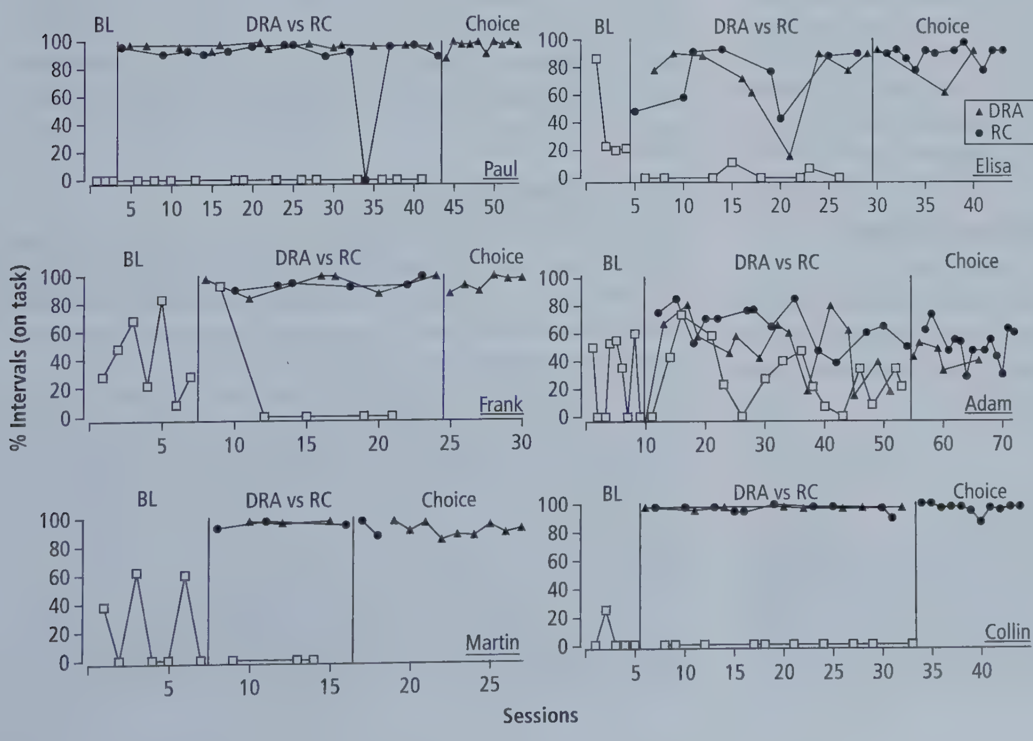


Figure 8.17 Multielement design with an initial baseline and a final participant's choice phase.

"Efficacy of and preference for reinforcement and response cost in token economies" by E. S. Jowett Hirst, C. L. Dozier, and S. W. Payne, 2016, *Journal of Applied Behavior Analysis*, *49*, p. 340. Reproduced with permission of John Wiley & Sons Inc.

three treatment procedures: FCT, FCT with punishment, and punishment only. Pressing one of the switches (the initial links in the chained schedule) resulted in a 2-minute period inside the therapy room (terminal links in the chain) during which the child experienced the contingencies corresponding to the switch he or she had selected. The FCT and FCT plus punishment conditions were carried out in the same manner as the treatment evaluation. When the punishment-only condition was selected, problem behavior resulted in a 30-seconds hands-down procedure for Jay or a 30-seconds hands-down and visual-screen procedure for Betty. Prior to using switch pressing to evaluate treatment preference, each child was prompted to push the different switches and was exposed to the different contingencies associated with each switch (four exposure sessions with Jay, two with Betty).

During treatment preference sessions, after each two-link chain (switch-press followed by 2 minutes of the selected contingencies for problem behavior), the child then exited the therapy room and was repositioned in front of the switches. The procedure was repeated until 20 minutes had elapsed for Jay; or until 10 initial-link responses were recorded for Betty. Hanley and colleagues (2005) described the results of the treatment preference evaluation (see Figure 8.18):

> Jay allocated the majority (79.3%) of his switch presses to the FCT plus punishment intervention. A preference for FCT plus punishment emerged during Betty's evaluation, in that all switch presses were allocated to FCT plus punishment in the final five sessions. These data provide direct evidence that the function-based treatments involving punishment contingencies were more preferred (reinforcing) than similar function-based treatments involving extinction or interventions consisting of punishment alone. (p. 61)

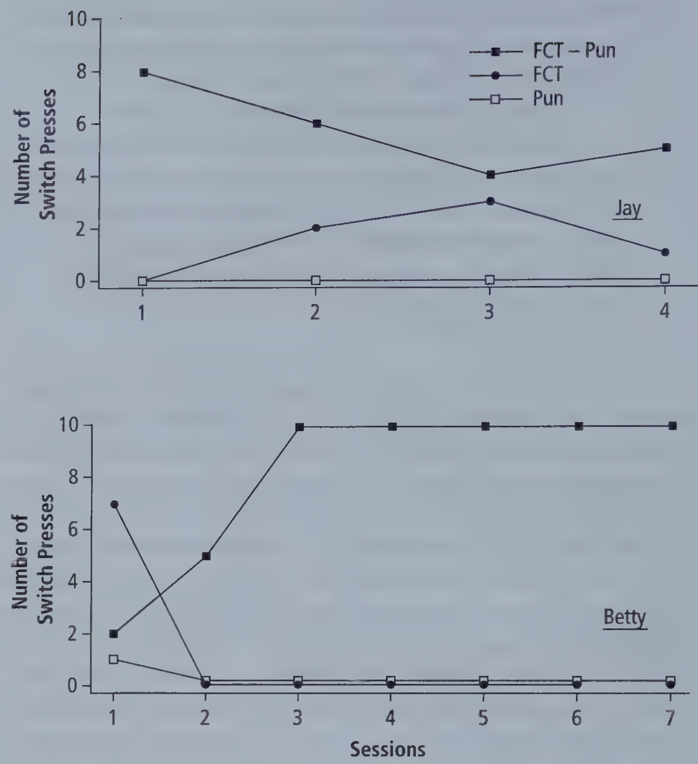


Figure 8.18 Concurrent chains design, participant's choice phase.

Based on "On the effectiveness of and preference for punishment and extinction components of function-based interventions" by G. P. Hanley, C. C. Piazza, W. W. Fisher, and K. A. Maglieri, 2005, *Journal of Applied Behavior Analysis*, *38*, p. 60. Copyright 2005 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

Readers should not view the variations of the multielement design described here as a catalog of possible and appropriate applications of this experimental tactic. In countless other ways, a multielement/alternating treatments tactic can be incorporated within an experimental design employing one or more reversal to baseline phases, multiple baselines, and changing criterion designs (e.g., Cariveau, Kodak, & Campbell, 2016; Coon & Miguel, 2012; Saini, Gregory, Uran, & Fantetti, 2015).

Advantages of the Multielement Design

The multielement design offers numerous advantages for evaluating and comparing two or more independent variables. Most of the benefits cited here were described by Ulman and Sulzer-Azaroff (1975), who first brought the rationale and possibilities of the multielement design to the attention of the applied behavior analysis community.

Treatment Withdrawal Not Required

A major advantage of the alternating treatments design is that it does not require the investigator to withdraw a seemingly effective treatment to demonstrate a functional relation. Reversing behavioral improvements raises ethical issues that can be avoided with the alternating treatments design. Regardless of ethical concerns, however, administrators and teachers may be more likely to accept a multielement design over a reversal design even when one of the alternating treatments is a no-treatment control condition. "It would appear that a return to baseline conditions every other day or every third day is not as disagreeable to a teacher as is first establishing a high level of desirable behavior for a prolonged period, and then reinstating the baseline behaviors" (Ulman & Sulzer-Azaroff, 1975, p. 385).

Allows Rapid Comparison

The experimental comparison of two or more treatments can often be made quickly with the multielement design. The ability of this design to rapidly demonstrate differential effects is a major reason it is the basic experimental tactic used in functional analysis of problem behavior (see Chapter 27).

When the effects of different treatments become apparent early in a multielement design, the investigator can then switch to programming only the most effective treatment. The efficiency of the multielement design can leave a researcher with meaningful data even when an experiment must be terminated early. A reversal or multiple baseline design, in contrast, must be carried through to completion to show a functional relation.

Researchers and practitioners must guard against judging that a functional relation exists between outcomes associated with two or more conditions in a multielement design when no

such relation exists. Such false positives are highly likely when only one or two measures are taken per condition. For example, Kahng and Iwata (1999) found that false positives were obtained more often in brief multielement designs than in multielement designs with three or more sessions in each condition.

Equally important is discovering ineffective treatments. For example, multielement designs enabled researchers to demonstrate that colored overlays did not improve reading fluency of children diagnosed with dyslexia (Freeze Denton & Meindle, 2016); and that weighted vests did not increase appropriate in-seat behaviors of elementary-age students with autism and severe to profound intellectual disabilities (Cox, Gast, Luscre, & Ayers, 2009).

Minimizes Irreversibility Problem

Some behaviors, even though they have been brought about or modified by an intervention, do not return to baseline levels when the intervention is withdrawn and thereby resist analysis with an A-B-A-B design. However, rapidly alternating treatment and no-treatment (baseline) conditions may reveal differences in responding between the two conditions, especially early in an experiment before responding in the no-treatment condition begins to approximate the level of responding in the treatment condition.

Minimizes Sequence Effects

A properly conducted multielement design minimizes the extent to which sequence effects confound the experiment's results. Sequence effects pose a threat to the internal validity of any experiment, but especially to studies involving multiple treatments. The concern over sequence effects can be summed up by this simple question: Would the results have been the same had the sequence of treatments been different? Sequence effects can be extremely difficult to control in experiments using reversal or multiple baseline tactics (see Chapter 9) to compare two or more independent variables because each experimental condition must remain in effect for a fairly long period, thereby producing a specific sequence of events. However, in a multielement design, independent variables can be rapidly alternated in a random fashion that produces no particular sequence. Also, each treatment is in effect for short periods, reducing the likelihood of carryover effects (O'Brien, 1968). The multielement design's ability to minimize sequence makes it a powerful tool for achieving complex behavior analyses.

Accommodates Unstable Data

Determining functional behavior–environment relations in the presence of unstable data presents a serious problem for the applied behavior analyst. Using steady state responding to predict, verify, and replicate behavioral changes is the foundation of experimental reasoning in behavior analysis (Sidman, 1960). However, obtaining stable baseline responding with many socially important behaviors of interest to applied behavior analysts is extremely difficult. Merely providing a subject with repeated opportunities to emit a target response can result in gradually improved performance. Although practice effects are worthy of empirical investigation because of their applied and scientific importance, the unstable baselines they create pose problems for the analysis of intervention variables. The changing levels of task difficulty inherent in moving through a curriculum of progressively more complex material also make obtaining steady state responding for many academic behaviors difficult.

Because the different treatment conditions are alternated rapidly in a multielement design, because each condition is presented many times throughout each time period encompassed by the study, and because no single condition is present for any considerable length of time, it can be presumed that any effects of practice, change in task difficulty, maturation, or other historical variables are equally represented in each treatment condition and therefore will not differentially affect any one condition more or less than the others. For example, even though each of two data paths representing a student's reading performance under two different teaching procedures shows variable and ascending trends that might be due to practice effects and uneven curriculum materials, any consistent separation and vertical distance between the data paths can be attributed to differences in the teaching procedures.

Reveals Generalization of Effects

By alternating various conditions of interest, an experimenter can continually assess the degree of generalization of behavior change from an effective treatment to other conditions of interest. For example, by alternating different therapists in the final phase of their study of pica behavior, Singh and Winton (1985) demonstrated the extent to which the overcorrection treatment was effective when presented by different persons.

Intervention Can Begin Immediately

Although determining the preintervention level of responding is generally preferred, the clinical necessity of immediately attempting to change some behaviors precludes repeated measurement in the absence of intervention. When necessary, a multielement design can begin without an initial baseline phase.

Considering the Appropriateness of the Multielement Design

The advantages of the multielement design are significant. As with any experimental tactic, however, the multielement design presents certain disadvantages and leaves unanswered certain questions that can be addressed only by additional experimentation.

Multiple Treatment Interference

The fundamental feature of the multielement design is the rapid alternation of two or more independent variables irrespective of the behavioral measures obtained under each treatment. Although the rapid alternation minimizes sequence effects and reduces the time required to compare treatments, it raises the important question of whether the effects observed under any of the alternated treatments would be the same if each treatment

were implemented alone. **Multiple treatment interference** refers to the confounding effects of one treatment on a subject's behavior being influenced by the effects of another treatment administered in the same study.

Multiple treatment interference must always be suspected in the multielement design. However, by following the alternating treatments phase with a phase in which only one treatment is in effect, the experimenter can assess the effects of that treatment when administered in isolation.

Unnatural Nature of Rapidly Alternating Treatments

The rapid back-and-forth switching of treatments does not reflect the typical manner in which clinical and educational interventions are applied. From an instructional perspective, rapid switching of treatments can be viewed as artificial and undesirable. In most instances, however, the quick comparison of treatments offered by the multielement design compensates for its contrived nature. The concern of whether participants might suffer detrimental effects from the rapid alternation of conditions is an empirical question that can be determined only by experimentation. Also, it is helpful for practitioners to remember that one purpose of the multielement design is to identify an effective intervention as quickly as possible so that the participant does not have to endure ineffective instructional approaches or treatments that would delay progress toward educational or clinical goals. On balance, the advantages of rapidly switching treatments to identify efficacious interventions will usually outweigh any undesirable effects that such manipulation may cause.

Limited Capacity

Although the multielement design enables an elegant, scientifically sound method for comparing the differential effects of two or more treatments, it is not an open-ended design with which an unlimited number of treatments can be compared. Although multielement designs with up to five conditions have been reported (e.g., Didden, Prinsen, & Sigafoos, 2000), in most situations a maximum of four different conditions (one of which may be a no-treatment control condition) can be compared effectively within a given phase of an alternating treatments design, and in many instances only two different treatments can be accommodated. To separate the effects of each treatment condition from effects that may be caused by aspects of the alternating treatments design, each treatment must be carefully counterbalanced across all potentially relevant aspects of its administration (e.g., time of day, order of presentation, settings, therapists). In many applied settings, the logistics of counterbalancing more than two or three treatments would be overly cumbersome and require too many sessions to complete the experiment. Also, too many competing treatments might hamper the subject's ability to discriminate between treatments, thereby reducing the design's effectiveness.

Selection of Treatments

Theoretically, although a multielement design can be used to compare the effects of any two discrete treatments, in reality the design is more limited. To enhance the probability of discrimination between conditions, and thereby the possibility of obtaining reliable, measurable differences in behavior, the treatments should embody significant differences from one to the other. For example, an investigator using a multielement design to study the effects of group size on students' academic performance during instruction might include conditions of 4-student, 10-student, and 20-student groups. Alternating conditions of 6-, 7-, and 8-student groups is less likely to reveal differences between group size and performance. However, a treatment condition should not be selected for inclusion in a multielement design only because it might yield a data path easily differentiated from that of another condition.

The *applied* in applied behavior analysis encompasses the nature of treatment conditions as well as the nature of behaviors investigated (Wolf, 1978). An important consideration in selecting treatment conditions should be the extent to which they represent current practices or practices that could conceivably be implemented. Although an experiment comparing the effects of 5, 10, and 20 minutes of math homework per school night on math achievement might be useful, a study comparing the effects of 5 minutes, 10 minutes, and 3 hours of math homework per night would not be. Even if such a study found 3 hours of nightly math homework effective in raising students' achievement in math, few teachers, parents, administrators, or students would carry out a program of 3 hours of nightly homework for a single content area.

Another consideration is that some interventions may not produce socially significant behavior change unless and until they have been implemented consistently for a continuous period of time.

> When a multielement baseline design is employed, overlapping data do not necessarily rule out the possible efficacy of an experimental procedure. The session-by-session alternation of conditions might obscure effects that could be observed if the same condition was presented during several consecutive sessions. It is therefore possible that a given treatment may prove to be effective with a reversal or multiple baseline design, but not with a multielement baseline design. (Ulman & Sulzer-Azaroff, 1975, p. 382)

The suspicion that a given treatment may be effective if presented in isolation for an extended period is an empirical question that can be explored through experimentation. At one level, if extended application of a single treatment results in behavioral improvement, the practitioner might be satisfied, and no further action would be needed. However, the practitioner-researcher who is interested in determining experimental control might return to a multielement design and compare the performance of the single treatment with that of another intervention.

SUMMARY

Reversal Design

1. The reversal tactic (A-B-A) entails repeated measures of behavior in a given setting during three consecutive phases: (a) baseline (independent variable absent), (b) treatment (independent variable present), and (c) return to baseline (independent variable absent).
2. Reintroducing the independent variable in the form of an A-B-A-B design strengthens the reversal design tremendously. The A-B-A-B design is the most straightforward and generally most powerful intrasubject design for demonstrating functional relations.

Variations of the A-B-A-B Design

3. Extending the A-B-A-B design with repeated reversals may provide a more convincing demonstration of a functional relation than a design with one reversal.
4. The B-A-B reversal design can be used with target behaviors for which an initial baseline phase is inappropriate or not possible for ethical or practical reasons.
5. Multiple treatment reversal designs (e.g., A-B-A-B-C-B-C, A-B-C-A-C-B-C) use the reversal tactic to compare the effects of two or more experimental conditions to baseline and/or to one another.
6. Multiple treatment reversal designs are particularly susceptible to confounding by sequence effects.
7. The noncontingent reinforcement (NCR) reversal technique enables the isolation and analysis of the contingent aspect of reinforcement.
8. Reversal techniques incorporating differential reinforcement control conditions (DRO and DRI/DRA) can also be used to analyze the effects of contingent reinforcement.

Considering the Appropriateness of the Reversal Design

9. Experimental designs featuring the reversal tactic are ineffective in evaluating the effects of treatment variables that, by their very nature, cannot be withdrawn once presented (e.g., instruction, modeling).
10. Once changed during treatment, some behaviors will not reverse to baseline levels when the independent variable has been withdrawn. Such behavioral irreversibility precludes effective use of the reversal design.
11. Withdrawing a seemingly effective treatment variable to provide scientific verification of its function in changing behavior may raise legitimate social, educational, and ethical concerns.
12. Brief reversal phases, sometimes even one-session baseline probes, may be sufficient to demonstrate believable experimental control.

Multielement Design

13. The multielement design (also called alternating treatments design) compares two or more treatments (i.e., independent variables) while their effects on the target behavior (i.e., dependent variable) are measured.
14. In a multielement design, each successive data point for a specific treatment plays three roles: (a) a basis for the *prediction* of future levels of responding under that treatment, (b) potential *verification* of the previous prediction of performance under that treatment, and (c) the opportunity for *replication* of previous effects produced by that treatment.
15. Experimental control is demonstrated in a multielement design when the data paths for two different treatments show little or no overlap.
16. The extent of differential effects produced by two treatments is determined by the vertical distance between their respective data paths and quantified by the vertical axis scale.

Variations of the Multielement Design

17. Common variations of the multielement (alternating treatments) design include the following:
    - Single-phase design alternating two or more conditions (one of which may be a no-treatment control condition)
    - Two-phase design: initial baseline and comparison phase
    - Three-phase design: initial baseline, alternating treatments, and best treatment only
    - Three-phase design: initial baseline, alternating treatments, and participant's choice
    - Concurrent chains (or schedule) design

Advantages of the Multielement Design

18. Advantages of the multielement (alternating treatments) design include the following:
    - Does not require treatment withdrawal.
    - Rapidly compares the relative effectiveness of treatments.

- Minimizes the irreversibility problem.
- Minimizes sequence effects.
- Can be used with unstable data patterns.
- Can assess generalization of effects.
- Intervention can begin immediately.

Considering the Appropriateness of the Multielement Design

19. Multielement designs are susceptible to multiple treatment interference. However, by following the alternating treatments phase with a phase in which only one treatment is administered, the experimenter can assess the effects of that treatment in isolation.
20. The rapid back-and-forth switching of treatment conditions does not typify clinical or education practice and may be viewed as artificial and undesirable.
21. An alternating treatments phase is usually limited to a maximum of four different treatment conditions.
22. Multielement designs are most effective in revealing the differential effects of treatment conditions that differ significantly from one another.
23. Multielement designs are not effective for assessing the effects of independent variables that change behavior only when implemented over a continuous period of time.

KEY TERMS

A-B-A design
A-B-A-B design
adapted alternating treatments design
alternating treatments design
B-A-B design
concurrent chains (schedule) design
DRI/DRA reversal technique
DRO reversal technique
irreversibility
multielement design
multiple treatment interference
multiple treatment reversal design
noncontingent reinforcement (NCR) reversal technique
reversal design
sequence effects
withdrawal design

MULTIPLE-CHOICE QUESTIONS

1. All things being equal, an experiment that incorporates ________ reversal(s) presents a more convincing and compelling demonstration of experimental control than an experiment with _______ reversal(s).
   - **a.** Multiple, one
   - **b.** One, four
   - **c.** Several, multiple
   - **d.** Five, seven

   Hint: (See "Reversal Design")
2. Which of the following is considered to be a reason for using a B-A-B design instead of an A-B-A-B design?
   - **a.** Irreversibility
   - **b.** Unlimited time in which to demonstrate practical and social significance
   - **c.** Sequence effects
   - **d.** When treatment is already in place

   Hint: (See "Reversal Design")
3. This design is considered the most "straightforward" and generally the most powerful within-subject design for demonstrating a functional relationship.
   - **a.** Alternating treatments design
   - **b.** Multiple baseline design
   - **c.** AB design
   - **d.** A-B-A-B reversal design

   Hint: (See "Reversal Design")
4. For most investigations, it is suggested that you only change _______ variable(s) at a time.
   - **a.** One
   - **b.** Two
   - **c.** Three
   - **d.** Four

   Hint: (See "Reversal Design")
5. _________ control conditions might be used as a variation of the reversal design when it is not possible or when it is deemed inappropriate to completely eliminate an event or activity as a reinforcer.
   - **a.** Difference
   - **b.** Noncontingent reinforcement
   - **c.** Sequenced
   - **d.** Temporal

   Hint: (See "Reversal Design")

6. This experimental tactic entails repeated measures of behavior in a given setting that require at least three consecutive phases: baseline, introduction of the independent variable, and a return to baseline.
   - a. Alternating treatments design
   - b. Multiple baseline design
   - c. Repeated multiple measures design
   - d. ABA reversal design

   Hint: (See "Reversal Design")

7. ________ means that a level of behavior observed in an earlier phase cannot be reproduced even though experimental conditions return to preintervention conditions.
   - a. Sequence effects
   - b. Irreversibility
   - c. Order effects
   - d. Reliability

   Hint: (See "Reversal Design")

8. This design is predicated on the behavioral principle of stimulus discrimination and allows for the comparison of two or more different treatments.
   - a. Multiple baseline design
   - b. Alternating treatments design
   - c. A-B-A-B reversal design
   - d. Single-subject design

   Hint: (See "Multielement Design")

9. An investigator who ________ turns the target behavior on and off by presenting and withdrawing a specific variable makes a clear and convincing demonstration of experimental control.
   - a. Conditionally
   - b. Haphazardly
   - c. Covertly
   - d. Reliably

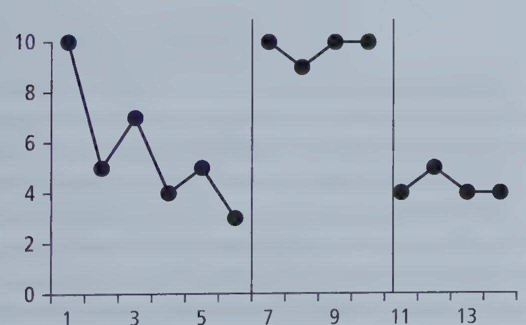
   Hint: (See "Reversal Design")

10. This term is used to describe the effects on a subject's behavior in a given condition that may be a result of the subject's experience with a prior condition.
    - a. Sequence
    - b. Irreversible
    - c. Temporal
    - d. Variable

    Hint: (See "Multielement Design")

ESSAY-TYPE QUESTIONS

1. Describe the experimental tactic shown in the figure. Discuss how the design allows for the detection of a functional relationship between the dependent and independent variables of interest.



   Hint: (See "Reversal Design")

2. You are interested in comparing the effects of two distinctly different schedules of reinforcement on students' academic engagement. Select an appropriate experimental tactic and describe its implementation.

   Hint: (See "Multielement Design")

3. State one disadvantage in the use of a reversal design and in the use of an alternating treatments design in applied settings.

   Hint: (See "Reversal Design" and "Multielement Design")

4. Summarize how the reversal design incorporates the elements of baseline logic (prediction, verification, and replication).

   Hint: (See "Multielement Design")

5. State one ethical consideration in the use of the reversal design as an experimental tactic in an applied school setting.

   Hint: (See "Multielement Design")

NOTES

1. Some authors use the term **withdrawal design** to describe experiments based on an A-B-A-B analysis and reserve the term *reversal design* for studies in which the behavioral focus of the treatment variable is reversed (or switched to another behavior), as in the DRO and DRI/DRA reversal techniques described later in this chapter. However, *reversal design*, as behavior analysts most often use the term, encompasses both withdrawals and reversals of the independent variable, signifying the investigator's attempt to demonstrate "behavioral reversibility" (Baer, Wolf, & Risley, 1968; Thompson & Iwata, 2005). Also, *withdrawal design* is sometimes used to describe an experiment in which the treatment variables are sequentially or partially withdrawn after their effects have been analyzed in an effort to promote maintenance of the target behavior (Rusch & Kazdin, 1981).
2. Risley (1997, 2005) credits Montrose Wolf with designing the first experiments using the reversal and multiple baseline designs. "The research methods that Wolf pioneered in these studies were groundbreaking. That methodology came to define applied behavior analysis" (pp. 280–281).
3. Strictly speaking, using NCR as an experimental control technique to demonstrate that the contingent application of reinforcement is requisite to its effectiveness is not a separate variation of the A-B-A reversal design. Technically, the NCR reversal technique, as well as the differential reinforcement of other behavior (DRO) and differential reinforcement of incompatible behavior (DRI)/differential reinforcement of alternative behavior (DRA) reversal techniques described next, is a multiple treatment design. For example, the Baer and Wolf (1970a) study of social reinforcement shown in Figure 8.6 used an A-B-C-B-C design, with B representing the NCR conditions and C representing the contingent reinforcement conditions.
4. The extent to which the increased diversity of the children's block building can be attributed to the attention and praise ("That's nice") or the descriptive feedback (". . . that's different") in the teacher's comments cannot be determined from this study because social attention and descriptive feedback were delivered as a package.
5. Additional manipulations in the form of the partial or sequential withdrawal of intervention components are made when it is necessary or desirable for the behavior to continue at its improved level in the absence of the complete intervention (cf., Rusch & Kazdin, 1981).
6. The mand is one of six types of elementary verbal operants identified by Skinner (1957). Chapter 18 describes Skinner's analysis of verbal behavior and its importance to applied behavior analysis.
7. Stimulus preference assessment procedures are described in Chapter 11.
8. DRA and response cost are described in Chapters 25 and 15, respectively.
9. Some published studies using a concurrent schedule design are described by their authors as using a *simultaneous treatment design*. Barlow and Hayes (1979), who introduced the term *simultaneous schedule design*, could find only one true example in the applied literature: a study by Browning (1967) in which three techniques for reducing the bragging of a 10-year-old boy were compared.

CHAPTER 9

Multiple Baseline and Changing Criterion Designs

LEARNING OBJECTIVES

- Describe the multiple baseline technique.
- Describe the differences between types of multiple baseline designs.
- Describe the changing criterion design technique.
- Explain how to systematically manipulate independent variables to analyze effects on treatment.
- Identify and explain the practical and ethical considerations in using various experimental designs.

This chapter describes two additional experimental tactics for analyzing behavior–environment relations: the multiple baseline design and the changing criterion design. In a multiple baseline design, after collecting initial baseline data simultaneously across two or more behaviors, settings, or people, the behavior analyst then applies the treatment variable sequentially across these behaviors, settings, or people and notes the effects. The changing criterion design is used to analyze improvements in behavior as a function of stepwise, incremental criterion changes in the level of responding required for reinforcement. In both designs, experimental control and a functional relation are demonstrated when the behaviors change from a steady state baseline to a new steady state after the introduction of the independent variable is applied, or a new criterion established.

MULTIPLE BASELINE DESIGN

The multiple baseline design is the most widely used experimental design for evaluating treatment effects in applied behavior analysis. It is a highly flexible tactic that enables researchers and practitioners to analyze the effects of an independent variable across multiple behaviors, settings, and/or subjects without having to withdraw the treatment variable to verify that the improvements in behavior were a direct result of the application of the treatment. As you recall from Chapter 8, the reversal design by its very nature requires that the independent variable be withdrawn to verify the prediction established in baseline. This is not so with the multiple baseline design.

Operation and Logic of the Multiple Baseline Design

Baer, Wolf, and Risley (1968) first described the **multiple baseline design** in the applied behavior analysis literature. They presented the multiple baseline design as an alternative to the reversal design for two situations: (a) when the target behavior is likely to be irreversible (e.g., once changed by the treatment variable, the behavior comes into contact with naturally existing contingencies of reinforcement or other sources of control that maintain the new level of performance) or (b) when it is undesirable, impractical, or unethical to withdraw a seemingly effective treatment condition. Additionally, some interventions, such as training or instructions, cannot be withdrawn. Terminating an ongoing instructional program does not re-establish conditions that existed prior to training; the participant "has been taught."

Figure 9.1 illustrates Baer and colleagues' explanation of the basic operation of the multiple baseline design.

> In the multiple baseline technique, a number of responses are identified and measured over time to provide baselines against which changes can be evaluated. With these baselines established, the experimenter then applies an experimental variable to one of the behaviors, produces a change in it, and perhaps notes little or no change in the other baselines. If so, rather than reversing the just-produced change, he instead applies the experimental variable to one of the other, as yet unchanged, responses. If it changes at that point, evidence is accruing that the experimental variable is indeed effective, and that the prior change was not simply a matter of coincidence. The variable then may be applied to still another response, and so on. The experimenter is attempting to show that he has a reliable experimental variable, in that each behavior changes maximally only when the experimental variable is applied to it. (p. 94)

The multiple baseline design takes three primary forms:

- The multiple baseline across behaviors design, consisting of two or more different behaviors of the same subject.
- The multiple baseline across settings design, consisting of the same behavior of the same subject in two or more different settings, situations, or time periods.
- The multiple baseline across subjects design, consisting of the same behavior of two or more different participants (or groups).

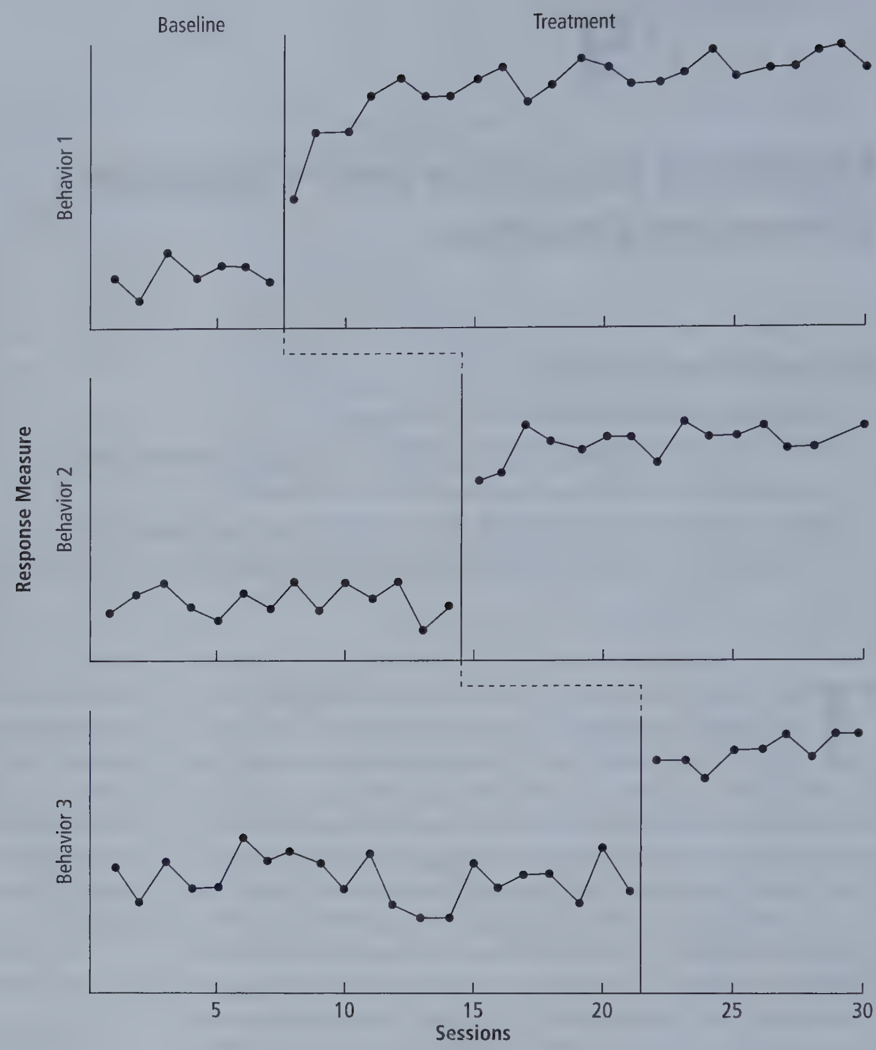


Figure 9.1 Graphic prototype of a multiple baseline design.

Although only one of the multiple baseline design's basic forms is called an "across behaviors" design, all multiple baseline designs involve the time-lagged application of a treatment variable across technically different (meaning independent) behaviors. That is, in the multiple baseline across settings design, even though the subject's performance of the same target behavior is measured in two or more settings, each behavior–setting combination is conceptualized and treated as a different behavior for analysis. Similarly, in a multiple baseline across subjects design, each subject–behavior combination functions as a different behavior in the operation of the design.

Figure 9.2 shows the same data set displayed in Figure 9.1, with the addition of data points representing predicted measures if baseline conditions were not changed and shaded areas illustrating how the three elements of baseline logic—prediction, verification, and replication—are operationalized in the multiple baseline design.[1] When stable baseline responding has been achieved for Behavior 1, a *prediction* is made that if the environment were held constant, continued measurement would reveal similar levels of responding. When the researcher's confidence in such a prediction is justifiably high, the independent variable is applied to Behavior 1. The open data points in the treatment phase for Behavior 1 represent the predicted level of responding. The solid data points show the actual measures obtained for Behavior 1 during the treatment condition. These data show a discrepancy with the predicted level of responding if no changes had been made in the environment, thereby suggesting that the treatment may be responsible for the change in behavior. The data collected for Behavior 1 in a multiple baseline design serve the same functions as the data collected during the first two phases of an A-B-A-B reversal design.

Continued baseline measures of the other behaviors in the experiment offer the possibility of verifying the prediction made for Behavior 1. In a multiple baseline design, *verification* of a predicted level of responding for one behavior (or tier) is obtained if little or no change is observed in the data paths of the behaviors (tiers) that are still exposed to the conditions under which the prediction was made. In Figure 9.2 those portions of the baseline condition data paths for Behaviors 2 and 3 within the shaded boxes verify the prediction for Behavior 1. At this point in the experiment, two inferences can be made: (a) The prediction that Behavior 1 would not change in a constant environment is valid because the environment was held constant for Behaviors 2 and 3 and their levels of responding remained unchanged; and (b) the observed changes in Behavior 1 were brought about by the independent variable because only Behavior 1 was exposed to the independent variable and only Behavior 1 changed.

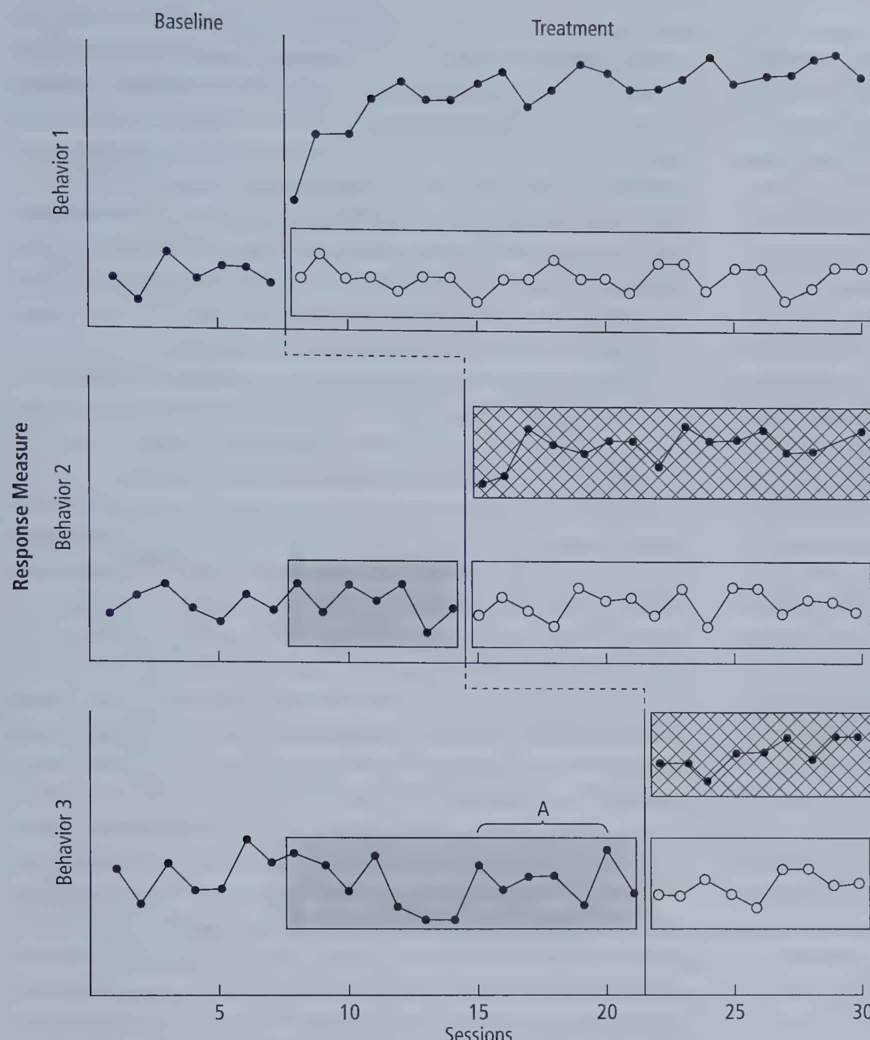


Figure 9.2 Graphic prototype of a multiple baseline design with shading added to show elements of baseline logic. Open data points represent predicted measures if baseline conditions were unchanged. Baseline data points for Behaviors 2 and 3 within shaded areas verify the prediction made for Behavior 1. Behavior 3 baseline data within Bracket A verify the prediction made for Behavior 2. Data obtained during the treatment condition for Behaviors 2 and 3 (cross-hatched shading) provide replications of the experimental effect.

In a multiple baseline design, the independent variable's function in changing a given behavior is inferred by the lack of change in untreated behaviors. However, verification of function is not demonstrated directly as it is with the reversal design, thereby making the multiple baseline design an inherently weaker tactic (i.e., less convincing from the perspective of experimental control) for revealing a functional relation between the independent variable and a target behavior. However, the multiple baseline design compensates somewhat for this weakness by providing the opportunity to verify or refute a series of similar predictions. Not only is the prediction for Behavior 1 in Figure 9.2 verified by continued stable baselines for Behaviors 2 and 3, but the bracketed portion of the baseline data for Behavior 3 also serves as verification of the prediction made for Behavior 2.

When the level of responding for Behavior 1 under the treatment condition has stabilized or reached a predetermined performance criterion, the independent variable is then applied to Behavior 2. If Behavior 2 changes in a manner similar to the changes observed for Behavior 1, *replication* of the independent variable's effect has been achieved (shown by the data path shaded with cross-hatching). After Behavior 2 has stabilized or reached a predetermined performance criterion, the independent variable is applied to Behavior 3 to see whether the effect will be replicated. The independent variable may be applied to additional behaviors in a similar manner until a convincing demonstration of the functional relation has been established (or rejected) and all of the behaviors targeted for improvement have received treatment.

As with verification, replication of the independent variable's specific effect on each behavior in a multiple baseline design is not manipulated directly. Instead, the generality of the independent variable's effect across the behaviors constituting the experiment is demonstrated by applying it to a series of behaviors. Assuming accurate measurement and proper experimental control of relevant variables (i.e., the only environmental factor that changes during the course of the experiment should be the presence—or absence—of the independent variable), each time a behavior changes when, and only when, the independent variable is introduced, confidence in the existence of a functional relation increases.

How many different behaviors, settings, or subjects must a multiple baseline design include to provide a believable demonstration of a functional relation? Baer, Wolf, and Risley (1968) suggested that the number of replications needed in any design is ultimately a matter to be decided by the consumers of the research. In this sense, an experiment using a multiple baseline design must contain the minimum number of replications necessary to convince those who will be asked to respond to the experiment and to the researcher's claims (e.g., teachers, administrators, parents, funding sources, journal editors). A two-tier

multiple baseline design is a complete experiment and can provide strong support for the effectiveness of the independent variable (e.g., Grossi & Heward, 1998 [see Figure 29.3]; Harding, Wacker, Berg, Rick, & Lee, 2004; Lindberg, Iwata, Roscoe, Worsdell, & Hanley, 2003 [see Figure 26.2]; McCord, Iwata, Galensky, Ellingson, & Thomson, 2001 [see Figure 6.6]; May, Downs, Marchant, & Dymond, 2016; Newstrom, McLaughlin, & Sweeney, 1999 [see Figure 28.13]; Pennington & McComas, 2017; Test, Spooner, Keul, & Grossi, 1990 [see Figure 23.10]). McClannahan, McGee, MacDuff, and Krantz (1990) conducted a multiple baseline design study in which the independent variable was sequentially implemented in an eight-tier design across 12 participants. Multiple baseline designs of three to five tiers are most common. When the effects of the independent variable are substantial and reliably replicated, a three- or four-tier multiple baseline design provides a convincing demonstration of experimental effect. Suffice it to say that the more replications one conducts, the more convincing the demonstration will be.

Some of the earliest examples of the multiple baseline design in the applied behavior analysis literature were studies by Risley and Hart (1968); Barrish, Saunders, and Wolf (1969); Barton, Guess, Garcia, and Baer (1970); Panyan, Boozer, and Morris (1970); and Schwarz and Hawkins (1970). Some of the pioneering applications of the multiple baseline technique are not readily apparent with casual examination: The authors may not have identified the experimental design as a multiple baseline design (e.g., Schwarz & Hawkins, 1970), and/or the now-common practice of stacking the tiers of a multiple baseline design one on the other so that all of the data can be displayed graphically in the same figure was not always used (e.g., Maloney & Hopkins, 1973; McAllister, Stachowiak, Baer, & Conderman, 1969; Schwarz & Hawkins, 1970).

In 1970, Vance Hall, Connie Cristler, Sharon Cranston, and Bonnie Tucker published a paper that described three experiments, each an example of one of the three basic forms of the multiple baseline design: across behaviors, across settings, and across subjects. Hall and colleagues' paper was important not only because it provided excellent illustrations that today still serve as models of the multiple baseline design, but also because the studies were carried out by teachers and parents, indicating that practitioners "can carry out important and significant studies in natural settings using resources available to them" (p. 255).

Multiple Baseline Across Behaviors Design

The **multiple baseline across behaviors design** begins with the concurrent measurement of two or more behaviors of each participant in the study. After steady state responding has been obtained under baseline conditions, the investigator applies the independent variable to one of the behaviors and maintains baseline conditions for the other behavior(s). When steady state or criterion-level performance has been reached for the first behavior, the independent variable is applied to the next behavior, and so on (e.g., Davenport, Alber-Morgan, Clancy, & Kranak, 2017; Downs, Miltenberger, Biedronski, & Witherspoon 2015; Gena, Krantz, McClannahan, & Poulson, 1996; Harding et al., 2004; Higgins, Williams, & McLaughlin, 2001 [see Figure 28.1]).

Ward and Carnes (2002) used a multiple baseline across behaviors design to evaluate the effects of self-set goals and public posting on the execution of three skills by five linebackers on a college football team: (a) *reads*, in which the linebacker positions himself to cover a specified area on the field on a pass play or from the line of scrimmage on a run; (b) *drops*, in which the linebacker moves to the correct position depending on the offensive team's alignment; and (c) *tackles*. A video camera recorded the players' movements during all practice sessions and games. Data were collected for the first 10 opportunities each player had with each skill. Reads and drops were recorded as correct if the player moved to the zone identified in the coaches' playbook; tackles were scored as correct if the offensive ball carrier was stopped.

Following baseline, each player met with one of the researchers, who described the player's mean baseline performance for a given skill. Players were asked to set a goal for their performances during practice sessions; no goals were set for games. The correct performances during baseline for all five players ranged from 60 to 80%, and all players set goals of 90% correct performance. The players were informed that their performance in each day's practice would be posted on a chart prior to the next practice session. A *Y* (yes) or an *N* (no) was placed next to each player's name to indicate whether he had met his goal. A player's performance was posted on the chart only for the skill(s) in intervention. The chart was mounted on a wall in the locker room where all players on the team could see it. The head coach explained the purpose of the chart to other players on the team. Players' performances during games were not posted on the chart.

The results for one of the players, John, are shown in Figure 9.3. John met or exceeded his goal of 90% correct performance during all practices for each of the three skills. Additionally, his improved performance generalized to games. The same pattern of results was obtained for each of the other four players in the study, illustrating that the multiple baseline across behaviors design is a single-subject experimental strategy in which each subject serves as his own control. Each player constituted a complete experiment, replicated in this case with four other participants.

Multiple Baseline Across Settings Design

In the **multiple baseline across settings design**, a single behavior of a person (or group) is targeted in two or more different settings or conditions (e.g., locations, times of day). After stable responding has been demonstrated under baseline conditions, the independent variable is introduced in one of the settings while baseline conditions remain in effect in the other settings. When maximum behavior change or criterion-level performance has been achieved in the first setting, the independent variable is applied in the second setting, and so on.

Roane, Kelly, and Fisher (2003) employed a multiple baseline across settings design to evaluate the effects of a treatment designed to reduce the rate at which an 8-year-old boy put inedible objects in his mouth. Jason, who had been diagnosed with autism, cerebral palsy, and moderate intellectual disabilities, had a history of putting objects such as toys, cloth, paper, tree bark, plants, and dirt into his mouth. Data on Jason's mouthing were obtained concurrently in a classroom, a playroom, and outdoors—three settings that contained a variety of inedible objects and where caretakers had reported Jason's mouthing to be problematic. Observers in each setting unobtrusively tallied the number of times Jason inserted an inedible object past the

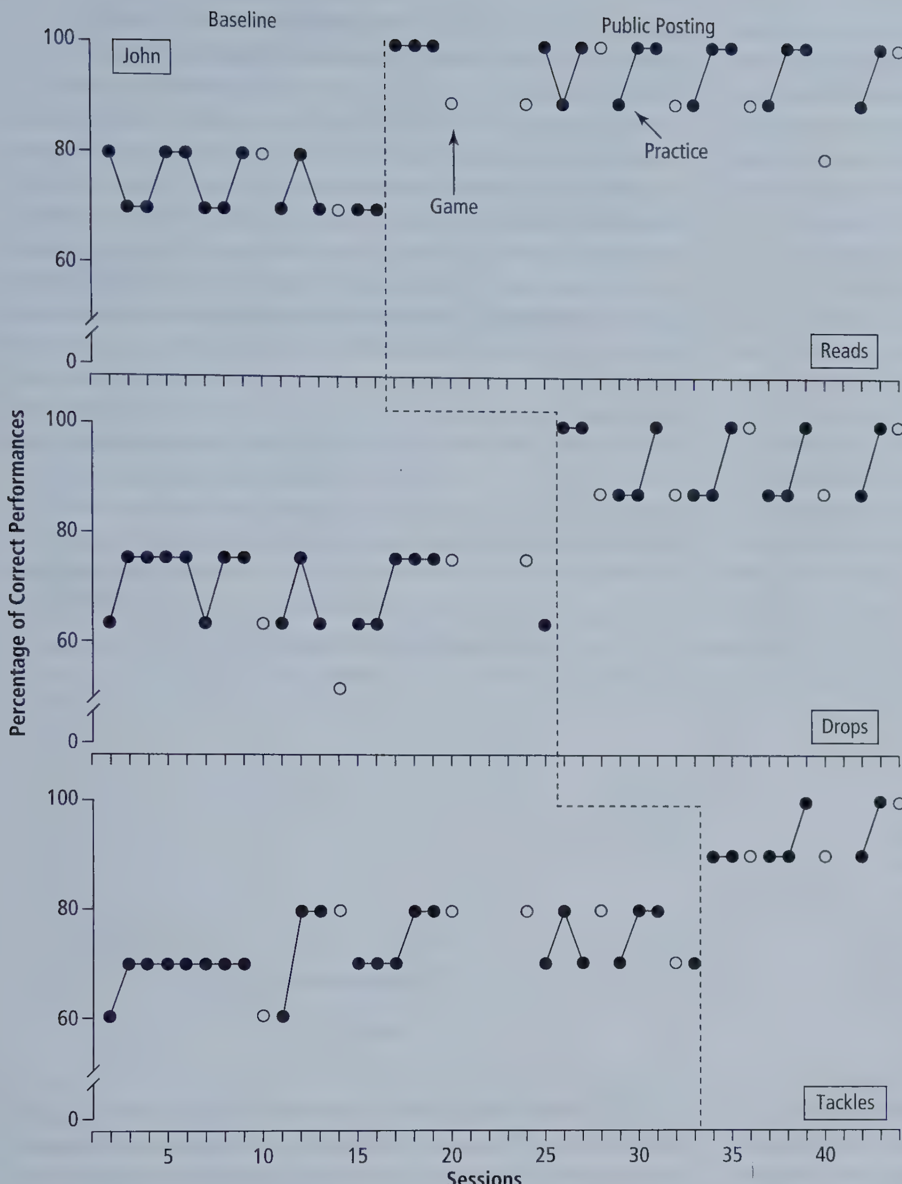


Figure 9.3 A multiple baseline across behaviors design showing percentage of correct reads, drops, and tackles by a college football player during practices and games.

Based on "Effects of Posting Self-Set Goals on Collegiate Football Players' Skill Execution During Practice and Games" by P. Ward and M. Carnes, 2002, *Journal of Applied Behavior Analysis, 35*, p. 5. Copyright 2002 by the Society for the Experimental Analysis of Behavior, Inc.

plane of his lips during 10-minute sessions. The researchers reported that Jason's object mouthing usually consisted of a series of discrete episodes, rather than an extended, continuous event, and that he often placed multiple objects (inedible objects and food) in his mouth simultaneously.

Roane and colleagues (2003) described the baseline and treatment conditions as follows:

> The baseline condition was developed based on the functional analysis results, which showed that mouthing was maintained by automatic reinforcement and occurred independent of social consequences. During baseline, a therapist was present (approximately 1.5 to 3 m from Jason), but all occurrences of mouthing were ignored (i.e., no social consequences were arranged for mouthing, and Jason was allowed to place items in his mouth). No food items were available during baseline. The treatment condition was identical to baseline except that Jason had continuous access to foods that had been previously identified to compete with the occurrence of object mouthing: chewing gum, marshmallows, and hard candy. Jason wore a fanny pack containing these items around his waist. (pp. 580–581)[2]

Figure 9.4 shows the staggered treatment sequence and results in each setting. During baseline, Jason mouthed objects at mean rates of 0.9, 1.1, and 1.2 responses per minute in the classroom, a playroom, and outdoor settings, respectively. Introduction of the fanny pack with food in each setting produced an immediate drop to a zero or near zero rate of mouthing. During treatment, Jason put items of food from the fanny pack into his mouth at mean rates of 0.01, 0.01, and 0.07 responses per minute in the classroom, a playroom, and outdoor settings, respectively. The multiple baseline across settings design revealed a functional relation between the treatment and a marked reduction in the frequency of Jason's object mouthing. No measures obtained during the treatment condition were as high as the lowest measures in baseline. During 22 of 27 treatment sessions across the three settings, Jason put no inedible objects in his mouth.

As was done in the Roane and colleagues' (2003) study, data for the different tiers in a multiple baseline across settings

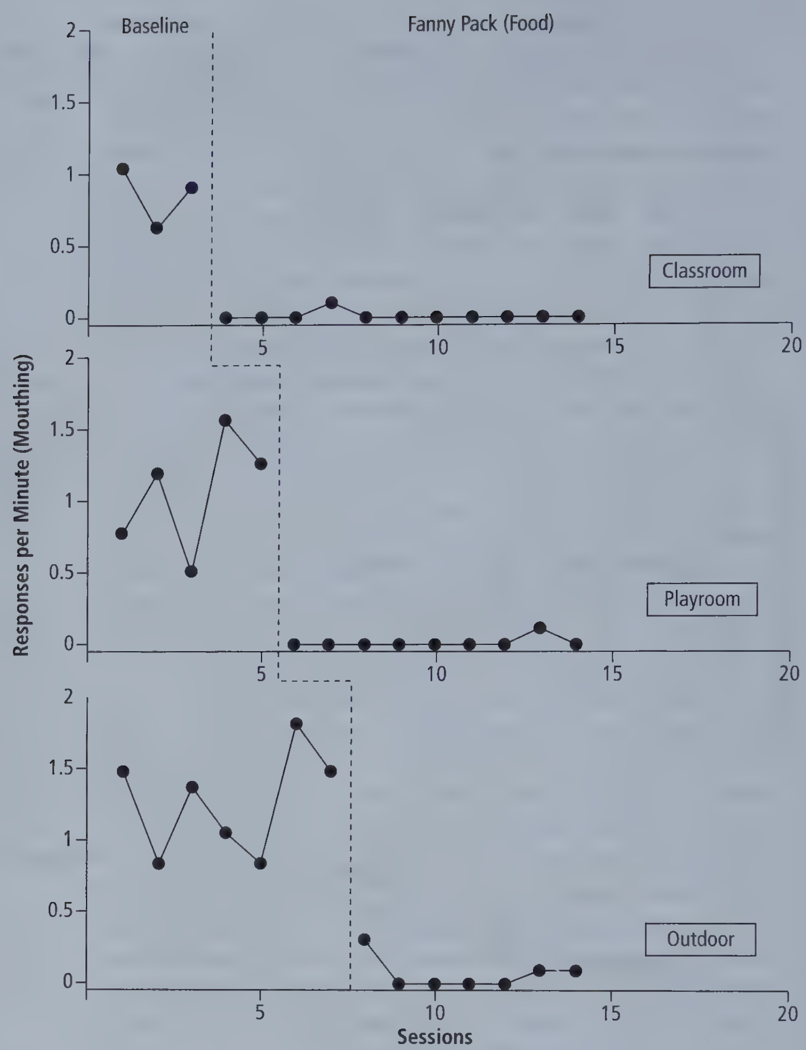


Figure 9.4 A multiple baseline across settings design showing the number of object mouthing responses per minute during baseline and treatment conditions.

Based on "The Effects of Noncontingent Access to Food on the Rate of Object Mouthing across Three Settings" by H. S. Roane, M. L. Kelly, and W. W. Fisher, 2003, *Journal of Applied Behavior Analysis, 36*, p. 581. Copyright 2003 by the Society for the Experimental Analysis of Behavior, Inc.

design are typically obtained in different physical environments (e.g., Cushing & Kennedy, 1997; Dalton, Martella, & Marchand-Martella, 1999). However, the different "settings" in a multiple baseline across settings design may exist in the same physical location and be differentiated from one another by different contingencies in effect, the presence or absence of certain people, and/or different times of the day. For example, in a study by Parker and colleagues (1984), the presence or absence of other people in the training room constituted the different settings (environments) in which the effects of the independent variable were evaluated. The attention, demand, and no-attention conditions (i.e., contingencies in effect) defined the different settings in a multiple baseline design study by Kennedy, Meyer, Knowles, and Shukla (2000; see Figure 6.4). The afternoon and the morning portions of the school day functioned as different settings in the multiple baseline across settings design used by Dunlap, Kern-Dunlap, Clarke, and Robbins (1991) to analyze the effects of curricular revisions on a student's disruptive and off-task behaviors.

In some studies using a multiple baseline across settings design, the participants are varied, changing, and perhaps even unknown to the researchers. For example, Van Houten and Malenfant (2004) used a multiple baseline design across two crosswalks on busy streets to evaluate the effects of an intensive driver enforcement program on the percentage of drivers yielding to pedestrians and the number of motor vehicle–pedestrian conflicts. Watson (1996) used a multiple baseline design across men's restrooms on a college campus to assess the effectiveness of posting signs in reducing bathroom graffiti.

Multiple Baseline Across Subjects Design

In the **multiple baseline across subjects design**, one target behavior is selected for two or more subjects in the same setting. After steady state responding has been achieved under baseline conditions, the independent variable is applied to one subject while baseline conditions remain in effect for the other subjects. When stable or criterion-level responding has been attained for the first subject, the independent variable is applied to another subject, and so on. The multiple baseline across subjects design is the most widely used of all three forms of the design, in part because practitioners are commonly confronted by multiple clients or students needing to learn the same skill or eliminate the same problem behavior (e.g., Craft, Alber, & Heward, 1998 [see Figure 30.10]; Kahng, Iwata, DeLeon, & Wallace, 2000 [see Figure 26.1]; Killu, Sainato, Davis, Ospelt, & Paul, 1998 [see Figure 26.3]; Kladopoulos & McComas, 2001 [see Figure 6.3]). Sometimes a multiple baseline design is conducted across "groups" of participants (e.g., Dixon & Holcomb, 2000 [see Figure 13.8]; Lewis, Powers, Kelk, & Newcomer, 2002 [see Figure 28.7]; White & Bailey, 1990 [see Figure 15.3]).

Scherrer and Wilder (2008) used a multiple baseline across subjects design to evaluate the effects of training of safe tray carrying by cocktail servers. Participants were three servers chosen randomly from the full-time staff of a bar, whose "managers reported that servers had complained of sore muscles and joints, particularly after a busy shift" (p. 132). The dependent variable was the percentage of eight safe tray-carrying behaviors (e.g., carrying position for amount of weight, finger and thumb positions, wrist position, tray distance from body) performed by each server when carrying trays to customers. During evening shifts, 3 to 4 days per week over the 8-week study, trained observers sitting at tables not assigned to the participants recorded safe tray-carrying behaviors on checklists they concealed. All three servers, when asked at the end of the study if they had been aware of data collectors recording their behavior, said they were unaware of the observers.

The independent variable was a single training session that began with a three-step procedure for each of the eight safe tray-carrying behaviors: (1) The trainer explained the correct position or technique, (2) the trainer modeled the correct position or technique, and (3) the participant described the correct position or technique and then demonstrated it. A series of trials were followed in which the server had to demonstrate each correct position or technique four times. All three participants met the mastery criteria in 30 to 50 minutes.

Figure 9.5 shows the results of the study. During baseline the three participants performed the safe carrying positions and techniques at means of 40% (Sara), 41% (Mike), and 49% (Tanya). Each participant's performance was variable, and no server attained as high as 70% safe performance on any baseline session. Although just three baseline measures were obtained for Sara, the sharp downward trend in the data indicated she should receive safety training first. Safe carrying behaviors by each participant increased immediately after training and remained at high and stable levels for the remainder of the study (means of 96% for Sara, 93% for Mike, and 96% for Tanya).

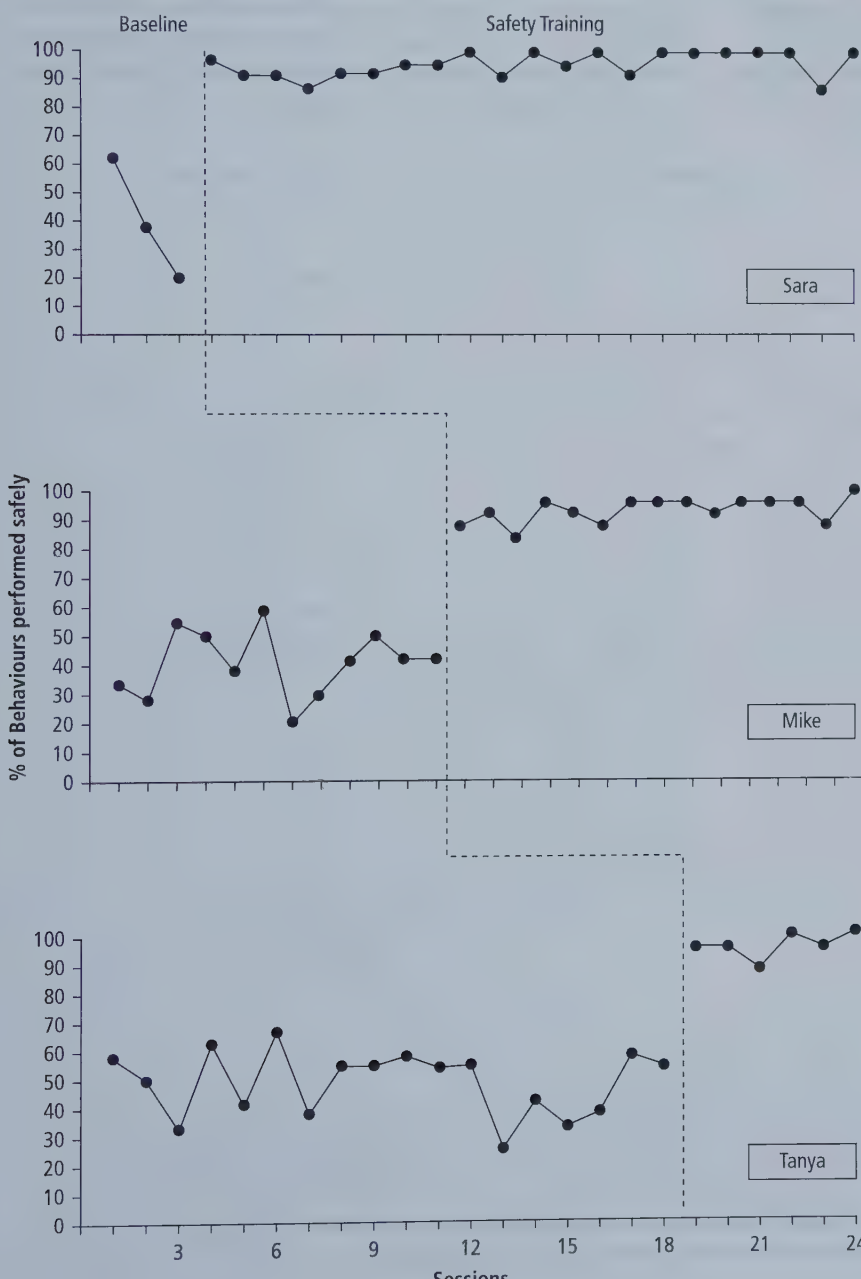


Figure 9.5 A multiple baseline across subjects design showing the percentage of safe tray-carrying positions and techniques by cocktail servers during baseline and post safety training conditions.

From "Training to Increase Safe Tray Carrying Among Cocktail Servers," by M. D. Scherrer and D. A. Wilder, 2008, *Journal of Applied Behavior Analysis, 41*, p. 134. Copyright 2008 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

The multiple baseline is not a true single-case design because each subject does not serve as his or her own control. Instead, verification of predictions based on the baseline data for each subject must be inferred from the relatively unchanging baseline measures of other subjects who are still in baseline, and replication of effects must be inferred from changes in the behavior of other subjects when they come into contact with the independent variable. This is both a weakness and a potential advantage of the multiple baseline across subjects design, and is discussed later in the chapter.

Variations of the Multiple Baseline Design

Three variations of the multiple baseline design are the multiple probe design, the delayed multiple baseline design, and the nonconcurrent multiple baseline across participants design.

Multiple Probe Design

The multiple probe design enables the behavior analyst to extend the operation and logic of the multiple baseline tactic to behaviors or situations in which concurrent measurement of all behaviors constituting the design is unnecessary, potentially reactive, impractical, or too costly. The **multiple probe design** was first described by Horner and Baer (1978) as a method for analyzing the relation between an independent variable and the acquisition of successive approximations or task sequence. In contrast to the standard multiple baseline design—in which data are collected simultaneously throughout the baseline phase for each behavior, setting, or subject in the experiment—in the multiple probe design, intermittent measures, or probes, provide the basis for determining whether behavior change has occurred prior to intervention. According to Horner and Baer, when applied to a chain or sequence of related behaviors to be learned, the multiple probe design provides answers to four questions: (a) What is the initial level of performance on each step (behavior) in the sequence? (b) What happens when sequential opportunities to perform each step in the sequence are provided prior to training on that step? (c) What happens to each step as training is applied? and (d) What happens to the performance of untrained steps in the sequence as criterion-level performance is reached on the preceding steps?

Figure 9.6 shows a graphic prototype of the multiple probe design. Although researchers have developed many variations of the multiple probe technique, the basic design has three key features: (a) An initial probe determines the subject's level of performance on each behavior in the sequence; (b) a series of baseline

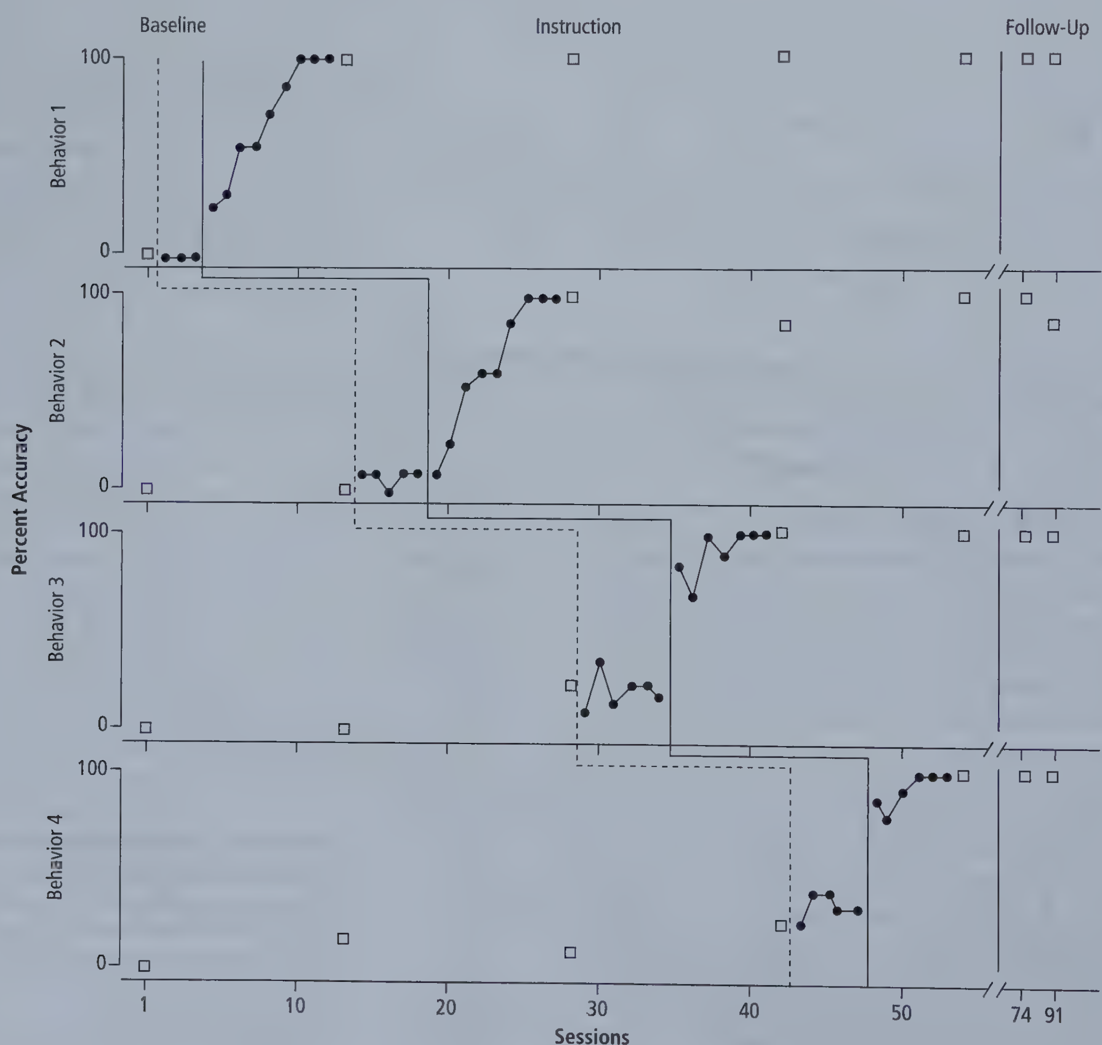


Figure 9.6 Graphic prototype of a multiple probe design. Square data points represent results of probe sessions in which the entire sequence or set of behaviors (1–4) is tested.

measures is obtained on each step prior to training on that step; and (c) after criterion-level performance is reached on any training step, a probe of each step in the sequence is obtained to determine whether performance has changed in any other steps.

Jimenez, Browder, Spooner, and Dibiase (2003) used a multiple probe design to analyze the effects of a peer-mediated instructional procedure on the acquisition of science knowledge by middle school students with moderate intellectual disability. Participants included five target students ages 11 to 14 with IQs ranging from 34 to 55 and six 11-year-old peer tutors from the same inclusive classroom of 26 students. Throughout the study, the general education teacher introduced each science inquiry lesson to the entire class, using the first three letters of a KWHL chart (K = what do you Know?; W = What do you want to know?; H = How will you find out?; L = what did you Learn?) to prompt student responses and thinking. Students then broke into groups of four or five, each group including one target student and one peer tutor, and participated in a hands-on experiment or online activity and reported what they saw.

The primary dependent variable was the number of correct responses by the target students on experimenter-administered assessment probes conducted after each science lesson. Each probe consisted of eight response opportunities: stating two printed science vocabulary words, naming two science pictures (e.g., student shown three flash cards with science vocabulary words or pictures and told, "find kinetic energy"), matching two word/picture combinations, and completing two concept statements (e.g., "Can you find the word that completes this statement? ______ is the energy of motion." [p. 305]). Three different picture sets were created for each vocabulary word. *In vivo* probes, conducted at least once per each unit of instruction, consisted of the target student's performance on the first peer-presented trial for each science response.

During intervention, the peer tutors (1) used a constant time-delay procedure to teach the target students to use the KWHL chart as the teacher led the class to complete their charts, and (2) embedded three constant time-delay learning trials for each science response during the small-group activity.[3] Figure 9.7 shows the results for one of the target students.

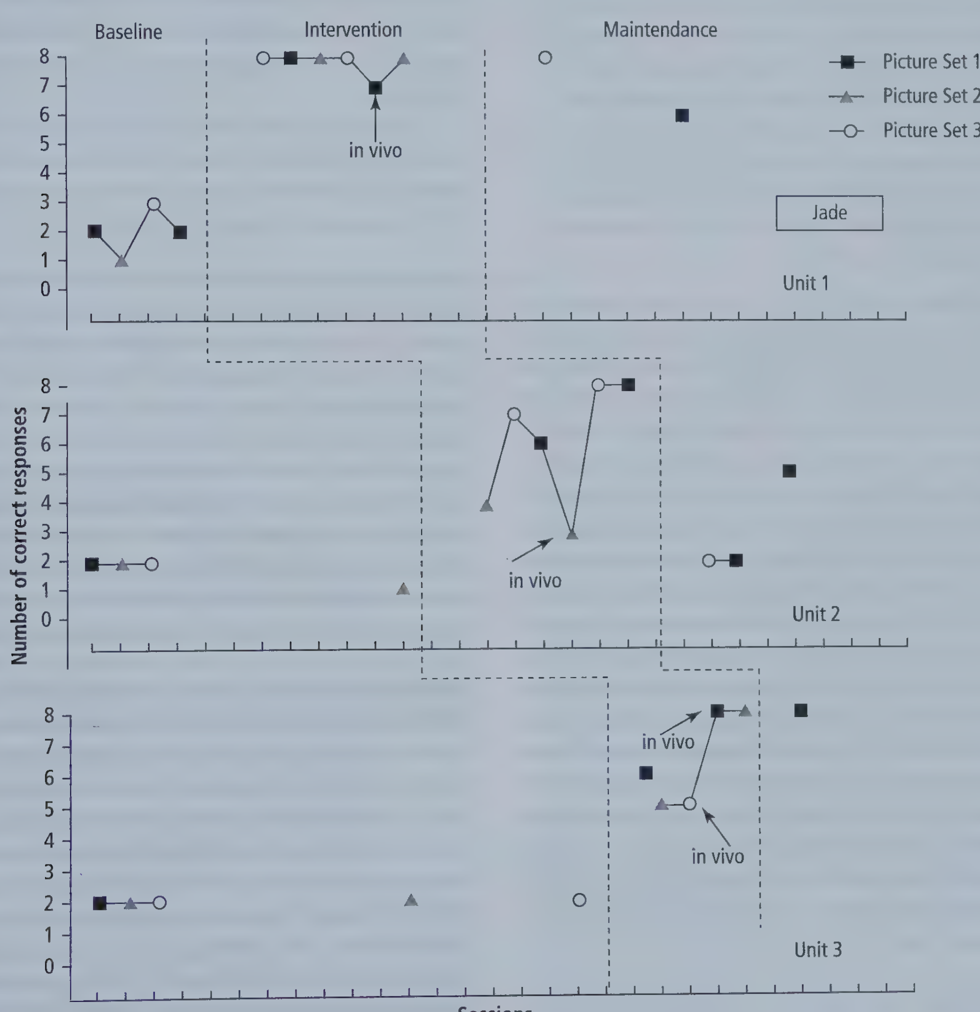


Figure 9.7 A multiple probe design across three curriculum units showing the number of correct responses on probes of science knowledge by a student with moderate intellectual disability as a function of peer-mediated embedded instruction.

From "Inclusive Inquiry Science Using Peer-Mediated Embedded Instruction for Students with Moderate Intellectual Disabilities," by B. A. Jimenez, D. M. Browder, F. Spooner, & D. Dibiase, 2012, *Exceptional Children, 78*, p. 310. Copyright 2012 by the Council for Exceptional Children. Reprinted by permission.

During baseline, Jade averaged 1.75 to 2 correct responses per probe across the three units. Her performance improved during intervention to a mean of 6 to 7.8 correct responses per probe. Jade's performance on maintenance probes remained well above baseline levels for two of the three units. The other four target students obtained similar results, suggesting a functional relation between peer-mediated embedded instruction and acquisition of science knowledge by students with moderate intellectual disability.

The multiple probe design is particularly appropriate for evaluating the effects of instruction on skill sequences in which it is highly unlikely that the subject can improve performance on later steps in the sequence without acquiring the prior steps. For example, the repeated measurement of a student's accuracy in solving division problems when the student possesses no skills in addition, subtraction, and multiplication would add little to an analysis. Horner and Baer (1978) made this point exceedingly well:

> The inevitable zero scores on the division baseline have no real meaning: division could be nothing else than zero (or chance, depending on the test format), and there is no real point in measuring it. Such measures are *pro forma:* they fill out the picture of a multiple baseline, true, but in an illusory way. They do not so much represent zero behavior as zero opportunity for the behavior to occur, and there is no need to document at the level of well-measured data that behavior does not occur when it cannot. (p. 190)

Thus, the multiple probe design avoids the necessity of ritualistic baseline data collection when the performance of any component of a chain or sequence is impossible or unlikely before acquisition of its preceding components. In addition to the two uses already mentioned—analysis of the effects of instruction on complex skill sequences and reduction in the amount of baseline measurement for behaviors that have no plausible opportunity to occur—the multiple probe technique is an effective experimental strategy for situations in which extended baseline measurement may prove reactive, impractical, or costly. The repeated measurement of a skill under nontreatment conditions can prove aversive to some students; and extinction, boredom, or other undesirable responses can occur. In his discussion of multiple baseline designs, Cuvo (1979) suggested that researchers should recognize that "there is a trade-off between repeatedly administering the dependent measure to establish a stable baseline on one hand and risking impaired performance by subjecting participants to a potentially punishing experience on the other hand" (pp. 222–223). Furthermore, complete assessment of all skills in a sequence may require too much time that could otherwise be spent on instruction.

Other examples of the multiple probe design can be found in Aldi et al. (2016; see Figure 21.2 in this text) Arntzen, Halstadtrø, and Halstadtrø (2003); Coleman-Martin and Wolff Heller (2004); Lambert and colleagues (2016); Mechling, Ayres, Purrazzella, and Purrazzella (2012); Werts, Caldwell, and Wolery (2016; see Figure 23.9 in this text) and Yanardag, Akmanoglu, and Yilmaz (2013).

Delayed Multiple Baseline Design

The delayed multiple baseline technique can be used when a planned reversal design is no longer possible or proves ineffective; it can also add additional tiers to an already operational multiple baseline design, as would be the case if new subjects were added to an ongoing study. The **delayed multiple baseline design** is an experimental tactic in which an initial baseline and intervention are begun, and subsequent baselines are added in a staggered or delayed fashion. Figure 9.8 shows a graphic prototype of the delayed multiple baseline design. The design employs the same experimental reasoning as a full-scale multiple baseline design, with the exception that data from baselines begun after the independent variable has been applied to previous behaviors, settings, or subjects cannot be used to verify predictions based on earlier tiers of the design. In Figure 9.8 baseline measurement of Behaviors 2 and 3 was begun early enough for those data to be used to verify the prediction made for Behavior 1. The final four baseline data points for Behavior 3 also verify the prediction for Behavior 2. However, baseline measurement of Behavior 4 began after the independent variable had been applied to each of the previous behaviors, thus limiting its role in the design to an additional demonstration of replication.

A delayed multiple baseline design may allow the behavior analyst to conduct research in certain environments in which other experimental tactics cannot be implemented. Heward (1978) suggested three such situations.

A Reversal Design Is No Longer Desirable or Possible. In applied settings the research environment may shift, negating the use of a previously planned reversal design. Such shifts may involve changes in the subject's environment that make the target behavior no longer likely to reverse to baseline levels, or changes in the behavior of parents, teachers, administrators, the subject/client, or the behavior analyst that, for any number of reasons, make a previously planned reversal design no longer desirable or possible. If other behaviors, settings, or subjects are appropriate for application of the independent variable, the behavior analyst could use a delayed multiple baseline technique and still pursue evidence of a functional relation.

Limited Resources, Ethical Concerns, or Practical Difficulties Preclude a Full-scale Multiple Baseline Design. This situation occurs when the behavior analyst only controls resources sufficient to initially record and intervene with one behavior, setting, or subject, and another research strategy is inappropriate. It may be that as a result of the first intervention, more resources become available for gathering additional baselines. This might occur following the improvement of certain behaviors whose pretreatment topography and/or rate required an inordinate expenditure of staff resources. Or, it could be that a reluctant administrator, after seeing the successful results of the first intervention, provides the resources necessary for additional analysis. Ethical concerns may preclude extended baseline measurement of some behaviors (e.g., Linscheid, Iwata, Ricketts, Williams, & Griffin, 1990). Also under this heading would fall the "practical difficulties" cited by Hobbs and Holt (1976) as a reason for delaying baseline measurement in one of three settings.

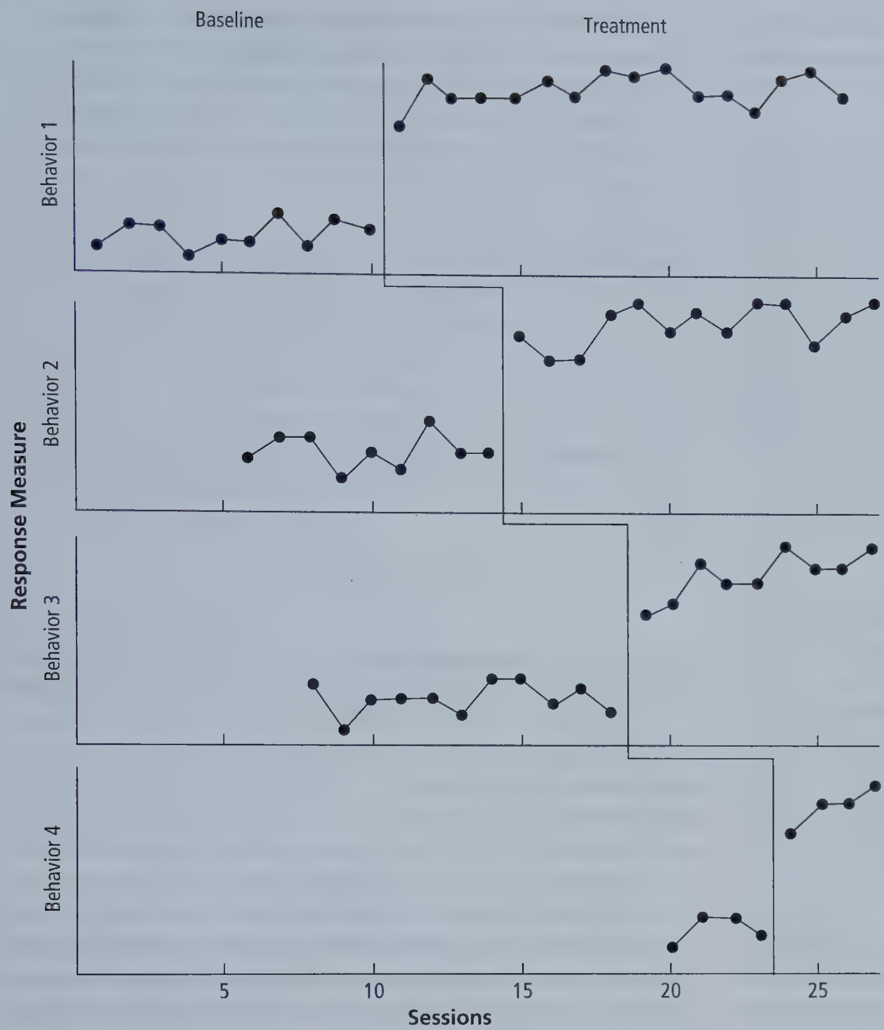


Figure 9.8 Graphic prototype of a delayed multiple baseline design.

"New" Behaviors, Settings, or Subjects Become Available. A delayed multiple baseline technique might become a useful analytic tactic due to changes in the environment (e.g., the subject begins to emit another behavior appropriate for intervention with the experimental variable, the subject begins to emit the original target behavior in another setting, or additional subjects displaying the same target behavior become available [e.g., Halldórsdóttir, Valdimarsdóttir, & Arntzen, 2017; Young & Daly, 2016, Figures 1 and 3]). Todd, Horner, and Sugai (1999) implemented a self-management intervention during Class Period B at the teacher's request because the teacher had noted an immediate and "dramatic change" in the student's performance when the intervention began in praise in Class Period A (see Figure 29.6).

Researchers have used delayed multiple baseline designs to evaluate a wide variety of interventions (e.g., Baer, Williams, Osnes, & Stokes, 1984; Jones, Fremouw, & Carples, 1977; Linscheid et al., 1990; Risley & Hart, 1968; Schepis, Reid, Behrmann, & Sutton, 1998; McKenzie, Smith, Simmons, & Soderlund, 2008, [see Figure 14.3]; Towery, Parsons, & Reid, 2014; White & Bailey, 1990 [see Figure 15.2]).[4] Poche, Brouwer, and Swearingen (1981) used a delayed multiple baseline design to evaluate the effects of a training program designed to prevent children from being abducted by adults. Three typically developing preschool children were selected as subjects because, during a screening test, each readily agreed to leave with an adult stranger. The dependent variable was the level of appropriateness of self-protective responses emitted by each child when an adult suspect approached the child and attempted to lure her away with a simple lure ("Would you like to go for a walk?"), an authoritative lure ("Your teacher said it was all right for you to come with me."), or an incentive lure ("I've got a nice surprise in my car. Would you like to come with me and see it?").

Each session began with the child's teacher bringing the child outdoors and then pretending to have to return to the building for some reason. The adult suspect (a confederate of the experimenters but unknown to the child) then approached the child and offered one of the lures. The confederate also served as observer, scoring the child's response on a 0 to 6 scale, with a score of 6 representing the desired response (saying, "No, I have to go ask my teacher" and moving at least 20 feet away from the suspect within 3 seconds) and a score of 0 indicating that the child moved some distance away from the school building with the suspect. Training consisted of modeling, behavioral rehearsal, and social reinforcement for correct responses.

Figure 9.9 shows the results of the training program. During baseline, all three children responded to the lures with

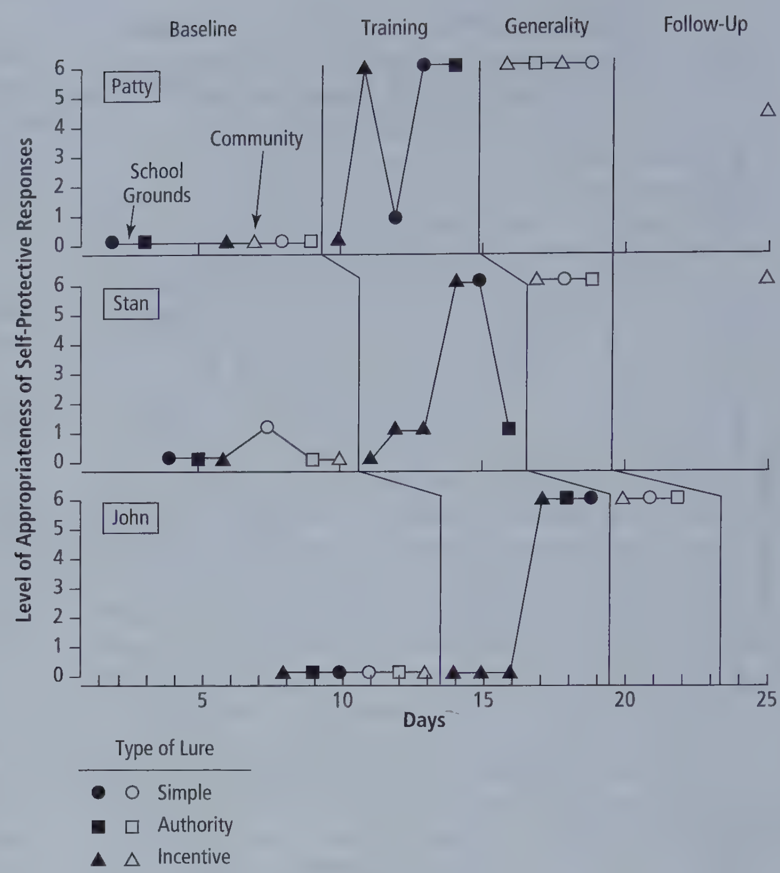


Figure 9.9 A delayed multiple baseline design showing the level of appropriateness of self-protective responses during baseline, training, and generality probes in school and community settings. Closed symbols indicate data gathered near the school; open symbols, in a location away from the school.

Based on "Teaching Self-Protection to Young Children," by C. Poche, R. Brouwer, and M. Swearingen, 1981, *Journal of Applied Behavior Analysis, 14*, p. 174. Copyright 1981 by the Society for the Experimental Analysis of Behavior, Inc.

safety ratings of 0 or 1. All three children mastered correct responses to the incentive lure in one to three training sessions, with one or two more sessions required for each child to master correct responses to the other two lures. Overall, training took approximately 90 minutes per child distributed over five or six sessions. All three children responded correctly when the lures were administered in generalization probes on sidewalk locations 150 to 400 feet from the school.

Although each baseline in this study was of equal length (i.e., had an equal number of data points), contradicting the general rule that the baselines in a multiple baseline design should vary significantly in length, there are two good reasons that Poche and colleagues began training when they did with each subject. First, the nearly total stability of the baseline performance of each child provided an ample basis for evaluating the training program (the only exception to complete susceptibility to the adult suspect's lures occurred when Stan stayed near the suspect instead of actually going away with him on his fourth baseline observation). Second, and more important, the nature of the target behavior required that it be taught to each child as soon as possible. Although continuing baseline measurement for varying lengths across the different tiers of any multiple baseline design is good practice from a purely experimental viewpoint, the ethics of such a practice in this instance would be highly questionable, given the potential danger of exposing the children to adult lures repeatedly while withholding training.

The delayed multiple baseline design entails several limitations. First, from an applied standpoint the design is not a good choice if the behavior analyst must wait too long to modify important behaviors, although this potential problem must be considered in all multiple baseline designs. Second, there is a tendency for the delayed baseline phases to contain fewer data points than found in a standard multiple baseline design, in which all baselines are begun simultaneously, resulting in baseline phases of considerable and varying length. Long baselines, if stable, provide the predictive power that permits convincing demonstrations of experimental control. Behavior analysts using any form of multiple baseline design must ensure that all baselines, regardless of when they begin, are of sufficient and varied length to provide believable bases for comparing experimental effects. A third limitation of the delayed multiple baseline design is that it can mask the interdependence of dependent variables.

> The strength of any multiple baseline design is that little or no change is noticed in the other, as yet untreated, behaviors until, and only until, the experimenter applies the independent variable. In a delayed multiple baseline design, the "delayed baseline" data gathered for subsequent behaviors may represent changed performance due to the experimental manipulation of other behaviors in the design and, therefore, may not be representative of the true, preexperimental operant level. . . . In such instances, the delayed multiple baseline might result in a "false negative," and the researcher may erroneously conclude that the intervention was not effective on the subsequent behavior(s), when in reality the lack of simultaneous baseline data did not permit the discovery that the behaviors covaried. This is a major weakness of the delayed

> multiple baseline design and makes it a research tactic of second choice whenever a full-scale multiple baseline can be employed. However, this limitation can and should be combated whenever possible by beginning subsequent baselines at least several sessions prior to intervention on previous baselines. (Heward, 1978, pp. 8–9)

Both the multiple probe design and the delayed multiple baseline design offer the applied behavior analyst alternative tactics for pursuing a multiple baseline analysis when extended baseline measurement is unnecessary, impractical, too costly, or unavailable. Perhaps the most useful application of the delayed multiple baseline technique is in adding tiers to an already operational multiple baseline design. Whenever a delayed baseline can be supplemented by probes taken earlier in the course of the study, experimental control is strengthened. As a general rule, the more baseline data, the better.

Nonconcurrent Multiple Baseline Across Participants Design

The **nonconcurrent multiple baseline across participants design** consists of a related series of A-B (baseline-intervention) sequences conducted across subjects at different points in time.

Watson and Workman (1981) proposed the nonconcurrent multiple baseline across participants design as a method for experimentally assessing treatment effects when concurrent measurement across subjects is not possible. They wrote, "The requirement [of a standard multiple baseline across subjects design] that observations be taken concurrently clearly poses problems for researchers in applied settings (e.g., schools, mental health centers), since clients with the same target behavior may only infrequently be referred at the same point in time" (p. 257, words in brackets added). The nonconcurrent multiple baseline across participants design is used frequently in applied behavior analysis (e.g., Garcia, Dukes, Brady, Scott, & Wilson, 2016; Greer, Neibert & Dozier, 2016 [see Figure 6.7]; Gunby, Carr, & LeBlanc, 2010; Kelley & Miltenberger, 2016; Speelman, Whiting, & Dixon, 2015).

"The nonconcurrent MB design is essentially a series of A-B replications in which the length of each baseline phase differs. The participant's graphs are then vertically aligned and visually inspected as one would a concurrent MB design" (Carr, 2005, p. 220). Although conjoining the graphs of two or more A-B designs with a doglegged dashed line produces a figure that "looks like" a multiple baseline design, doing so may suggest a greater degree of experimental control than is warranted (Harris & Jenson, 1985).

The nonconcurrent multiple baseline across participants design is inherently weaker than other multiple baseline design variations. Although the design entails two of the three elements of baseline logic—prediction and replication—the absence of concurrent baseline measures precludes the verification of intervention effects. We recommend labeling and scaling the horizontal axes of graphs of nonconcurrent multiple baseline design studies in a manner clearly depicting the actual time periods when data were obtained for each A-B sequence with respect to the others (e.g., see Carr, 2005; Harvey, May, & Kennedy, 2004, Figure 2).

Assumptions and Guidelines for Using Multiple Baseline Designs

Like all experimental tactics, the multiple baseline design entails certain assumptions about how the behavior–environment relations under investigation function, even though discovering the existence and operation of those relations is the very reason for conducting the research. In this sense, the design of behavioral experiments resembles an empirical guessing game—the experimenter guesses; the data answer. The investigator makes assumptions, hypotheses in the informal sense, about behavior and its relation to controlling variables and then constructs experiments designed to produce data capable of verifying or refuting those conjectures.[5]

Because verification and replication in the multiple baseline design depend on what happens, or does not happen, to other behaviors as a result of the sequential application of the independent variable, the experimenter must be particularly careful to plan and carry out the design in a manner that will afford the greatest degree of confidence in any relations suggested by the data. Although the multiple baseline design appears deceptively simple, its successful application requires much more than selecting two or more behaviors, settings, or subjects, collecting some baseline data, and then introducing a treatment condition to one behavior after the other. We suggest the following guidelines for designing and conducting experiments using multiple baseline designs.

Select Independent, yet Functionally Similar, Baselines

Demonstration of a functional relation in a multiple baseline design depends on two occurrences: (a) the behavior(s) still in baseline showing no significant change in level, trend, or variability while the behavior(s) in contact with the independent variable changes; and (b) each behavior changes when, and only when, the independent variable has been applied to it. Thus, the experimenter must make two, at times seemingly contradictory, assumptions about the behaviors targeted for analysis in a multiple baseline design. The assumptions are that the behaviors are functionally independent of one another (the behaviors will not covary with one another), and yet the behaviors share enough similarity that each will change when the same independent variable is applied to it. An error in either assumption can result in a failure to demonstrate a functional relation.

For example, let us suppose that the independent variable is introduced with the first behavior, and changes in level and/or trend are noted, but the other behaviors still in baseline also change. Do the changes in the still-in-baseline behaviors mean that an uncontrolled variable is responsible for the changes in all of the behaviors and that the independent variable is an ineffective treatment? Or do the simultaneous changes in the untreated behaviors mean that the changes in the first behavior were affected by the independent variable and have generalized to the other behaviors? Or, let us suppose instead that the first behavior changes when the independent variable is introduced, but subsequent behaviors do not change when the independent variable is applied. Does this failure to replicate mean that a factor other than the independent variable was responsible for

the change observed in the first behavior? Or does it mean only that the subsequent behaviors do not operate as a function of the experimental variable, leaving open the possibility that the change noted in the first behavior was affected by the independent variable?

Answers to these questions can and should be pursued by further experimental manipulations. In both kinds of failure to demonstrate experimental control, the multiple baseline design does not rule out the possibility of a functional relation between the independent variable and the behavior(s) that did change when the variable was applied. In the first instance, the failure to demonstrate experimental control with the originally planned design is offset by the opportunity to investigate and possibly isolate the variable robust enough to change multiple behaviors simultaneously. Discovery of variables that reliably produce generalized changes across behaviors, settings, and/or subjects is a major goal of applied behavior analysis; and if the experimenter is confident all other relevant variables were held constant before, during, and after the observed behavior changes, the original independent variable is the first candidate for further investigation.

In the second situation, failure to replicate changes from one behavior to another, the experimenter can pursue the possibility of a functional relation between the independent variable and the first behavior, perhaps using a reversal technique, and seek to discover later an effective intervention for the behavior(s) that did not change. Another possibility is to drop the original independent variable altogether and search for another treatment that might be effective with all of the targeted behaviors.

Select Concurrent and Plausibly Related Multiple Baselines

In an effort to ensure the functional independence of behaviors in a multiple baseline design, experimenters should not select response classes or settings so unrelated to one another as to offer no plausible means of comparison. For the ongoing baseline measurement of one behavior to provide the strongest basis for verifying the prediction of another behavior that has been exposed to an independent variable, two conditions must be met: (a) The two behaviors must be measured concurrently, and (b) all of the relevant variables that influence one behavior must have an opportunity to influence the other behavior. Studies that employ a multiple baseline approach across subjects and settings often stretch the logic of the design beyond its capabilities. For example, using the stable baseline measures of one child's compliance with parental requests as the basis for verifying the effect of intervention on the compliance behavior of another child living with another family is questionable practice. The sets of variables influencing the two children are surely differentiated by more than the presence or absence of the experimental variable.

> There are some important limits to designating multiple behavior/setting combinations that are intended to function as part of the same experiment. In order for the use of multiple behaviors and settings to be part of the same design and thus augment experimental reasoning, the general experimental conditions under which the two responses (whether two from one subject or one from each of two subjects) are emitted and measured must be ongoing concurrently. . . . Exposure [to the independent variable] does not have to be simultaneous for the different behavior/setting combinations, [but] it must be the identical treatment conditions along with the associated extraneous variables that impinge on the two responses and/or settings. This is because the conditions imposed on one behavior/setting combination must have the *opportunity* of influencing the other behavior/setting combination at the same time, regardless of the condition that actually prevails for the second. . . . It follows that using responses of two subjects each responding in different settings would not meet the requirement that there be a coincident opportunity for detecting the treatment effect. A treatment condition [as well as the myriad other variables possibly responsible for changes in the behavior of one subject] could not then come into contact with the responding of the other subject, because the second subject's responding would be occurring in an entirely different location. . . . Generally, the greater the plausibility that the two responses would be affected by the single treatment [and all other relevant variables], the more powerful is the demonstration of experimental control evidenced by data showing a change in only one behavior. (Johnston & Pennypacker, 1980, pp. 276–278)

The requirements of concurrency and plausible influence must be met for the verification element of baseline logic to operate in a multiple baseline design. However, replication of effect is demonstrated each time a baseline steady state is changed by the introduction of the independent variable, more or less regardless of where or when the variable is applied. Such nonconcurrent and/or unrelated baselines can provide valuable data on the generality of a treatment's effectiveness. See the previous discussion of the *nonconcurrent multiple baseline design*.

This discussion should not be interpreted to mean that a valid (i.e., logically complete) multiple baseline design cannot be conducted across different subjects, each responding in different settings. Numerous studies using mixed multiple baselines across subjects, response classes, and/or settings have contributed to the development of an effective technology of behavior change (e.g., Dixon et al., 1998; Durand, 1999 [see Figure 26.4]; Ryan, Ormond, Imwold, & Rotunda, 2002).

Let us consider an experiment designed to analyze the effects of a particular teacher training intervention, perhaps a workshop on using tactics to increase each student's opportunity to respond during group instruction. Concurrent measurement is begun on the frequency of student response opportunities in the classrooms of the teachers who are participating in the study. After stable baselines have been established, the workshop is presented first to one teacher (or group of teachers) and eventually, in staggered multiple baseline fashion, to all of the teachers.

In this example, even though the different subjects (teachers) are all behaving in different environments (different classrooms), comparison of their baseline conditions is experimentally sound because the variables likely to influence their teaching styles operate in the larger, shared environment in

which they all behave (the school and teaching community). Nevertheless, whenever experiments are proposed or published that involve different subjects responding in different settings, researchers and consumers should view the baseline comparisons with a critical eye toward their logical relation to one other.

Do Not Apply the Independent Variable to the Next Behavior Too Soon

To reiterate, for verification to occur in a multiple baseline design, it must be established clearly that as the independent variable is applied to one behavior and change is noted, little or no change is observed in the other, as-yet-untreated behaviors. The potential for a powerful demonstration of experimental control has been destroyed in many studies because the independent variable was applied to subsequent behaviors too soon. Although the operational requirement of sequential application in the multiple baseline tactic is met by introduction of the independent variable even in adjacent time intervals, the experimental reasoning afforded by such closely spaced manipulations is minimal.

> The influence of unknown, concomitant, extraneous variables that might be present could still be substantial, even a day or two later. This problem can be avoided by demonstrating continued stability in responding for the second behavior/setting combination during and after the introduction of the treatment for the first combination until a sufficient period of time has elapsed to detect any effect on the second combination that might appear. (Johnston & Pennypacker, 1980, p. 283)

Vary Significantly the Lengths of Multiple Baselines

Generally, the more the baseline phases in a multiple baseline design differ in the number of measures from one another, the stronger the design will be. Baselines with significantly different numbers of data points allow the unambiguous conclusion (assuming an effective treatment variable) that each behavior not only changes when the independent variable is applied, but also that each behavior does not change until the independent variable has been applied. If the different baselines are of the same or similar length, the possibility exists that changes noted when the independent variable is introduced are the result of a confounding variable, such as practice or reactivity to observation and measurement, and not a function of the experimental variable.

> Those effects . . . called practice, adaptation, warm-up, self-analysis, etc.; whatever they may be and whatever they may be called, the multiple baseline design controls for them by systematically varying the length of time (sessions, days, weeks) in which they occur prior to the introduction of the training package. . . . Such control is essential, and when the design consists of only two baselines, then the number of data points in each prior to experimental intervention should differ as radically as possible, at least by a factor of 2. I cannot see not systematically varying lengths of baselines prior to intervention, and varying them as much as possible/practical. Failure to do that . . . weakens the design too much for credibility. (D. M. Baer, personal communication June 2, 1978)

Intervene on the Most Stable Baseline First

In an ideal multiple baseline design, the independent variable is not applied to any of the behaviors until steady state responding has been achieved for each. However, applied behavior analysts are often denied the option of delaying treatment just to increase the strength of an experimental analysis. When intervention must begin before stability is evident across each tier of the design, the independent variable should be applied to the behavior, setting, or subject showing the most stable baseline data path (or counter-therapeutic trend, as with Sara in Figure 9.5). For example, if a study is designed to evaluate the effects of a teaching procedure on the rate of math computation of four students and no *a priori* reason exists for teaching the students in any particular sequence, instruction should begin with the student showing the most stable baseline. However, this recommendation should be followed only when the majority of the baselines in the design show reasonable stability.

Sequential applications of the independent variable should be made in the order of greatest stability at the time of each subsequent application. Again, however, the realities of the applied world must be heeded. The social significance of changing a particular behavior must sometimes take precedence over the desire to meet the requirements of experimental design.

Considering the Appropriateness of Multiple Baseline Designs

The multiple baseline design offers significant advantages, which no doubt have accounted for its widespread use by researchers and practitioners. Those advantages, however, must be weighed against the limitations and weaknesses of the design to determine its appropriateness in any given situation.

Advantages of the Multiple Baseline Design

Probably the most important advantage of the multiple baseline design is that it does not require withdrawing a seemingly effective treatment to demonstrate experimental control. This is a critical consideration for target behaviors that are self-injurious or dangerous to others. This feature of the multiple baseline design also makes it an appropriate method for evaluating the effects of independent variables that cannot, by their nature, be withdrawn and for investigating target behaviors that are likely or that prove to be irreversible (e.g., Duker & van Lent, 1991). Additionally, because the multiple baseline design does not necessitate a reversal of treatment gains to baseline levels, parents, teachers, or administrators may accept it more readily as a method of demonstrating the effects of an intervention.

The requirement of the multiple baseline design to sequentially apply the independent variable across multiple behaviors, settings, or subjects complements the usual practice of many practitioners whose goal is to develop multiple behavior changes. Teachers are charged with helping multiple students learn multiple skills to be used in multiple settings. Likewise, clinicians typically need to help their clients improve more than

one response class and emit more adaptive behaviors in several settings. The multiple baseline design is ideally suited to the evaluation of the progressive, multiple behavior changes sought by many practitioners in applied settings.

Because the multiple baseline design involves concurrent measurement of two or more behaviors, settings, or subjects, it is useful in assessing the occurrence of generalization of behavior change. The simultaneous monitoring of several behaviors gives the behavior analyst the opportunity to determine their covariation as a result of manipulations of the independent variable. Although changes in behaviors still under baseline conditions eliminate the ability of the multiple baseline design to demonstrate experimental control, such changes reveal the possibility that the independent variable is capable of producing behavioral improvements with desirable generality, thereby suggesting an additional set of research questions and analytic tactics (e.g., Odom, Hoyson, Jamieson, & Strain, 1985).

Finally, the multiple baseline design has the advantage of being relatively easy to conceptualize, thereby offering an effective experimental tactic for teachers and parents who are not trained formally in research methodology.

Limitations of the Multiple Baseline Design

The multiple baseline design presents at least three scientific limitations or considerations. First, a multiple baseline design may not allow a demonstration of experimental control even though a functional relation exists between the independent variable and the behaviors to which it is applied. Changes in behaviors still under baseline conditions and similar to concurrent changes in a behavior in the treatment condition preclude the demonstration of a functional relation within the original design. Second, from one perspective, the multiple baseline design is a weaker method for showing experimental control than the reversal design. This is because verification of the baseline prediction made for each behavior within a multiple baseline design is not directly demonstrated with that behavior, but must be inferred from the lack of change in other behaviors. This weakness of the multiple baseline design, however, should be weighed against the design's advantage of providing multiple replications across different behaviors, settings, or subjects. Third, the multiple baseline design provides more information about the effectiveness of the treatment variable than it does about the function of any particular target behavior.

> Consistently [the] multiple baseline is less an experimental analysis of the response than of the technique used to alter the response. In the reversal design, the response is made to work again and again; in the multiple-baseline design, it is primarily the technique that works again and again, and the responses either work once each [if different responses are used] or else a single response works once each per setting or once each per subject. Repetitive working of the same response in the same subject or the same setting is not displayed. But, while repetitive working of the response is foregone, repetitive and diverse working of the experimental technique is maximized, as it would not be in the reversal design. (Baer, 1975, p. 22)

Two important applied considerations that must be evaluated in determining the appropriateness of the multiple baseline design are the time and resources required for its implementation. First, because the treatment variable cannot be applied to subsequent behaviors, settings, or subjects until its effects have been observed on previous behaviors, settings, or subjects, the multiple baseline design requires that intervention be withheld for some behaviors, settings, or subjects, perhaps for a long time. This delay raises practical and ethical concerns. Treatment cannot be delayed for some behaviors; their importance makes delaying treatment impractical. And as Stolz (1978) pointed out, "If the intervention is generally acknowledged to be effective, denying it simply to achieve a multiple-baseline design might be unethical" (p. 33). Second, the resources needed for the concurrent measurement of multiple behaviors must be considered. Use of a multiple baseline design can be particularly costly when behavior must be observed and measured in several settings. However, when the use of intermittent probes during baseline can be justified in lieu of continuous measurement (Horner & Baer, 1978), the cost of concurrently measuring multiple behaviors can be reduced.

CHANGING CRITERION DESIGN

The changing criterion design can be used to evaluate the effects of a treatment that is applied in a graduated or stepwise fashion to a single target behavior. The changing criterion design was first described in the applied behavior analysis literature in two papers coauthored by Vance Hall (Hall & Fox, 1977; Hartmann & Hall, 1976).

Operation and Logic of the Changing Criterion Design

The reader can refer to Figure 9.10 before and after reading Hartmann and Hall's (1976) description of the **changing criterion design**.

> The design requires initial baseline observations on a single target behavior. This baseline phase is followed by implementation of a treatment program in each of a series of treatment phases. Each treatment phase is associated with a step-wise change in criterion rate for the target behavior. Thus, each phase of the design provides a baseline for the following phase. When the rate of the target behavior changes with each stepwise change in the criterion, therapeutic change is replicated and experimental control is demonstrated. (p. 527)

The operation of two elements of baseline logic—prediction and replication—is clear in the changing criterion design. When stable responding is attained within each phase of the design, a prediction of future responding is made. Replication occurs each time the level of behavior changes in a systematic way when the criterion is changed. Verification of the predictions based on each phase is not so obvious in this design but can be approached in two ways. First, varying the lengths of phases systematically enables a form of self-evident verification. The prediction is made that the level of responding will not change if the criterion is not changed. When the criterion is not changed

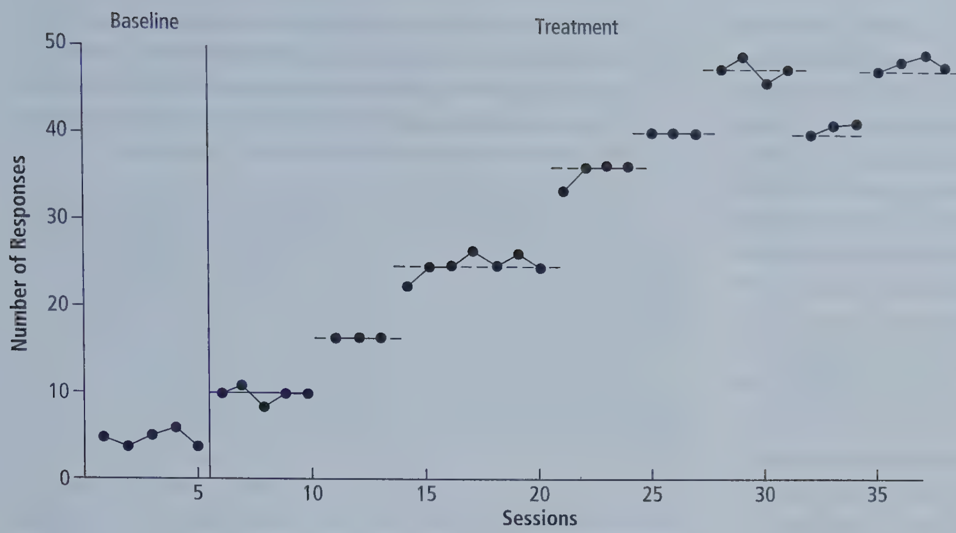


Figure 9.10 Graphic prototype of a changing criterion design.

and stable responding continues, the prediction is verified. When it can be shown within the design that levels of responding do not change unless the criterion is changed, regardless of the varied lengths of phases, experimental control is evident. Hall and Fox (1977) suggested another possibility for verification: "The experimenter may return to a former criterion and if the behavior conforms to this criterion level there is also a cogent argument for a high degree of behavioral control" (p. 154). Such a reversed criterion is shown in the next-to-last phase of Figure 9.10. Although returning to an earlier criterion level requires a brief interruption of the steady improvement in behavior, the reversal tactic strengthens the analysis considerably and should be included in the changing criterion design unless other factors indicate its inappropriateness.

One way to conceptualize the changing criterion design is as a variation of the multiple baseline design. Both Hartmann and Hall (1976, p. 530) and Hall and Fox (1977, p. 164) replotted data from changing criterion design experiments in a multiple baseline format, with each tier of the multiple baseline showing the occurrence or nonoccurrence of the target behavior at one of the criterion levels used in the experiment. A vertical condition change line doglegs through the tiers, indicating when the criterion for reinforcement was raised to the level represented by each tier. By graphing whether the target behavior was emitted during each session at or above the level represented on each tier both before and after the change in criterion to that level, a kind of multiple baseline analysis is revealed. However, the strength of the multiple baseline argument is not convincing because the "different" behaviors represented by each tier are not independent of one another. For example, if a target behavior is emitted 10 times in a given session, all of the tiers representing criteria below 10 responses would have to show that the behavior occurred, and all of the tiers representing criteria of 11 or more would have to show no occurrence of the behavior, or zero responding. The majority of the tiers that would appear to show verification and replication of effect, in fact, could only show these results because of the events plotted on another tier. A multiple baseline design provides a convincing demonstration of experimental control because the measures obtained for each behavior in the design are a function of the controlling variables for that behavior, not artifacts of the measurement of another behavior. Thus, recasting the data from a changing criterion design into a many-tiered multiple baseline format will often result in a biased picture in favor of experimental control.

Even though the multiple baseline design is not completely analogous, the changing criterion design can be conceptualized as a method of analyzing the increasing (or decreasing) performance of a target behavior. As Sidman (1960) pointed out, "It is possible to make reinforcement contingent upon a specified value of some aspect of behavior, and to treat that value as a response class in its own right" (p. 391). The changing criterion design can be an effective tactic for showing the repeated production of new rates of behavior as a function of manipulations of the independent variable (i.e., criterion changes).

Cook, Rapp, and Schulze (2015) used a changing criterion design to evaluate the effects of differential negative reinforcement of other behavior (DNRO) on the duration with which an 8-year-old boy with autism wore a medical alert bracelet. Jonah's parents were worried about their son's history of getting separated from caregivers in public places and purchased a medical alert bracelet that contained their contact information and Jonah's diagnosis.[6] Placing the bracelet on Jonah's wrist evoked problem behavior in the form of yelling, demanding the bracelet be removed, flopping on the floor, hitting people, biting his hand, and trying to remove the bracelet by pulling on it and shaking his wrist. Baseline consisted of 122 escape trials over 4 days. Each trial began with a therapist saying, "Jonah, it's time to wear your bracelet now," and placing the bracelet on Jonah's wrist. Contingent on problem behavior, the therapist immediately removed the bracelet for 30 seconds. When the escape period ended, the next trial began.

Each trial during the DNRO condition began with a therapist telling Jonah, "It's time to wear your bracelet now. If you can wait, I will take it off" (p. 903). The therapist placed the bracelet on Jonah's wrist and set a timer for the predetermined interval. If Jonah complied, the therapist simultaneously provided brief praise and removed the bracelet for the escape period. When the escape period ended, the next trial began. If Jonah was noncompliant, the therapist placed her hand over the bracelet to block his attempts to

remove it, described what Jonah needed to do to remove the bracelet (e.g., "Try to wait again with no yelling"), and reset the time to the current criterion wait time. The DNRO intervention began with a 5-second criterion for compliance time. Each time Jonah met criterion for five consecutive trials, the criterion was raised.

Jonah's mean latency to noncompliance during the escape baseline condition was 6 seconds (range 0 to 75 seconds) (see Figure 9.11; only the last 10 escape trials are plotted on the graph). Cook and colleagues (2015) described the sequence and magnitude of criterion changes and results:

> We introduced DNRO with a 5-s criterion and gradually increased the criterion to 60 s, which required 98 trials over 3 days. We then implemented DNRO with a 2-min criterion and extended the criterion to 45 min (middle panel); this increase required 64 trials over 5 days. We extended the duration that Jonah wore the bracelet from 1 hr to 7 hr in 37 trials over 12 days (bottom panel). Subsequently, therapists extended the intervention to Jonah's home and school (data available from the second author).
>
> Thereafter, apart from brief periods to allow for cleaning, Jonah reportedly wore the bracelet continuously for the next 2 years. (p. 905)

This study illustrates very well the changing criterion design's flexibility: the 21 criterion changes of magnitudes ranged from 5 seconds to 4 hours. Although one or more reversals to a previously attained criterion level may have provided a more convincing demonstration of experimental control (a limitation noted by the authors), the practical and ethical considerations of doing so would be questionable. As always, the applied behavior analyst must balance experimental concerns with the need to improve behavior in the most effective, efficient, ethical manner.

McDougall (2005) proposed a **range-bound changing criterion design** in which each intervention sub-phase includes both a lower and an upper criterion within which the participant

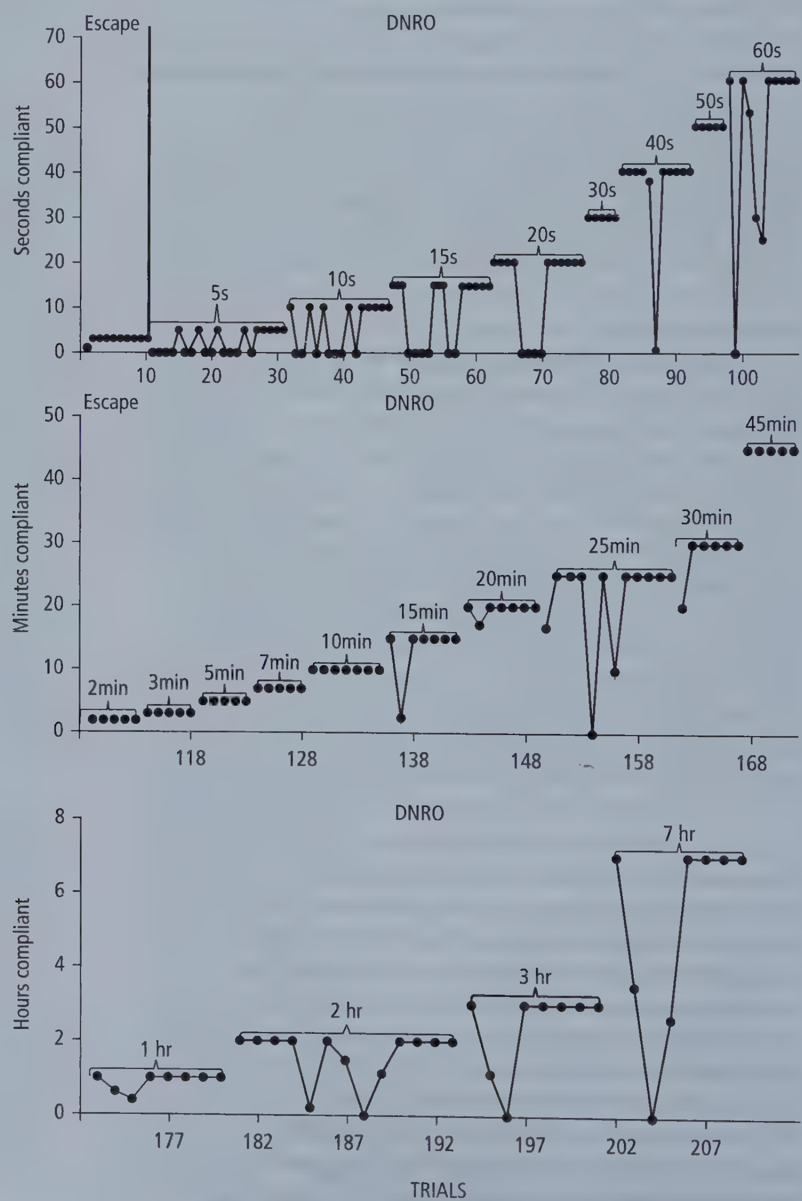


Figure 9.11 A changing criterion design showing how long an 8-year-old boy with autism complied with wearing a medical alert bracelet during baseline and differential negative reinforcement of other behavior (DNRO).

From "Differential Negative Reinforcement of Other Behavior to Increase Wearing of a Medical Bracelet," by J. L. Cook, J. T. Rapp, & K. A. Schulze, 2015, *Journal of Applied Behavior Analysis, 48,* p. 904. Copyright 2015 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

is expected to perform. McDougall suggested that behavioral "[i]mprovements as well as demonstrations of experimental control, are demonstrated definitively when the target behavior resides within, or conforms to, the pre-specified range" (p. 130). The author reported a study illustrating the design in which an overweight adult applied goal setting and self-graphing to increase minutes running each day (see Figure 9.12). After a 19-week baseline in which the participant ran on just 3 days, he established an initial aim of running an average of 20 minutes a day for 6 of 7 days of the week. He also created upper and lower ranges of ±10% of 20 minutes to serve as dual criteria for the minimum (18) and maximum (22) number of minutes he should run on any given day. When this aim was met, the participant increased his aim to 40 ± 4 minutes per day, then to 60 ± 6 minutes per day, and so on to a maximum of 100 ± 20 minutes per day, with one sub-phase reversing to lower criteria.

Compared to other single-case designs used in applied behavior analysis, relatively few studies using pure changing criterion designs have been reported (e.g., Allen & Evans, 2001; DeLuca & Holborn, 1992 [see Figure 13.2]; Foxx & Rubinoff, 1979; Kurti & Dallery, 2013; Wilson & Gratz, 2016). Some researchers have employed a changing criterion tactic as an analytic element within a larger experimental design (e.g., Martella, Leonard, Marchand-Martella, & Agran, 1993; Schleien, Wehman, & Kiernan, 1981).

Guidelines for Using the Changing Criterion Design

Proper implementation of the changing criterion design requires the careful manipulation of three design factors: length of phases, magnitude of criterion changes, and number of criterion changes. A survey of behavioral research literature published between 1971 and 2013 yielded 106 articles in 57 different journals reporting studies incorporating a changing criterion design (Klein, Houlihan, Vincent, & Panahon, 2017). The authors reported that just 25% of the graphs they reviewed met all necessary requirements for adequate experimental control. "Specifically, most studies lacked variation in phase magnitude (and did not compensate with a reversal), lacked variation in phase length, or implemented two or fewer changes in criterion; while some restricted responding so that criterions were forcibly met and or could not exceed the criterion" (p. 60).

Length of Phases

Because each phase in the changing criterion design serves as a baseline for comparing changes in responding measured in the next phase, each phase must be long enough to achieve stable responding. "Each treatment phase must be long enough to allow the rate of the target behavior to restabilize at a new and changed rate; it is stability after change has been achieved, and before introduction of the next change in criterion, that is crucial to producing a convincing demonstration of control" (Hartmann & Hall, 1976, p. 531). Target behaviors that are slower to change therefore require longer phases.

The length of phases in a changing criterion design should vary considerably to increase the design's validity. For experimental control to be evident in a changing criterion design, the target behavior not only must change to the level required by each new criterion in a predictable (preferably immediate) fashion, but also must conform to the new criterion for as long as it is in effect. When the target behavior

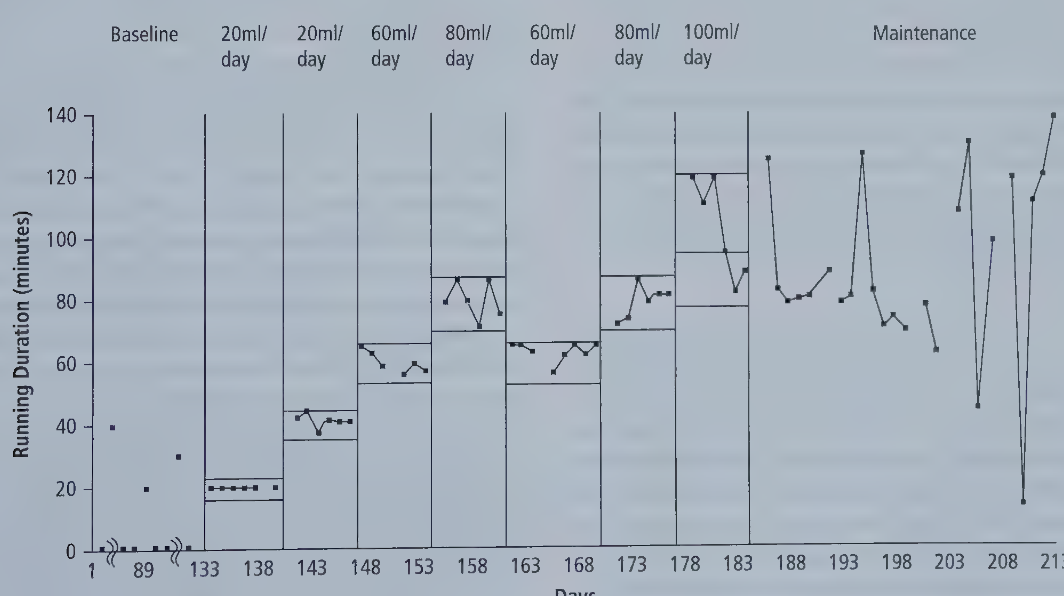


Figure 9.12 Example of a range-bound changing criterion design.

From "The Range-bound Changing Criterion Design," by D. McDougall, 2005, *Behavioral Interventions, 20,* p. 132. Copyright 2005 by John Wiley & Sons. Reprinted by permission.

closely follows successively more demanding criteria that are held in place for varied periods, the likelihood is reduced that the observed changes in behavior are a function of factors other than the independent variable (e.g., maturation, practice effects). In most situations, the investigator should not set a predetermined number of sessions for which each criterion level will remain in effect. It is best to let the data guide ongoing decisions, whether to extend the length of a current criterion phase or introduce a new criterion.

Magnitude of Criterion Changes

Varying the size of the criterion changes enables a more convincing demonstration of experimental control. When changes in the target behavior occur not only at the time a new criterion is implemented but also to the level specified by the new criterion, the probability of a functional relation is strengthened. In general, a target behavior's immediate change to meet a large criterion change is more impressive than a behavior change in response to a small criterion change. However, two problems arise if criterion changes are too large. First, setting aside practical considerations, and speaking from an experimental design standpoint only, large criterion changes may not permit inclusion of a sufficient number of changes in the design (the third design factor) because the terminal level of performance is reached sooner. The second problem is from an applied view: Criterion changes cannot be so large that they conflict with good instructional practice. Criterion changes must be large enough to be detectable, but not so large as to be unachievable. Therefore, the variability of the data in each phase must be considered in determining the size of criterion changes. Smaller criterion changes can be employed with very stable levels of responding, whereas larger criterion changes are required to demonstrate behavior change in the presence of variability.

When using a changing criterion design, behavior analysts must guard against imposing artificial ceilings (or floors) on the levels of responding that are possible in each phase. An obvious mistake of this sort would be to give a student only five math problems to complete when the criterion for reinforcement is five. Although the student could complete fewer than five problems, the possibility of exceeding the criterion has been eliminated, resulting perhaps in an impressive-looking graph, but one that is badly affected by poor experimental procedure.

Number of Criterion Changes

As a general rule, the more times the target behavior changes to meet new criteria, the more convincing the demonstration of experimental control. For example, eight criterion changes, one of which was a reversal to a previous level, were implemented in the changing design illustrated in Figure 9.10, and Cook et al. (2015) introduced 22 criterion changes in their study shown in Figure 9.11. Grey, Healy, Leader, and Hayes (2009) implemented 28 criterion changes in a study lasting 122 sessions. In both of these cases, a sufficient number of criterion changes occurred to demonstrate experimental control. The experimenter cannot, however, simply add any desired number of criterion changes to the design. Kratochwill and colleagues (2017) recommended a minimum of three criterion changes. The number of criterion changes that are possible within a changing criterion design is interrelated with the length of phases and the magnitude of criterion changes. Longer phases mean that the time necessary to complete the analysis increases; with a limited time to complete the study, the greater the number of phases, the shorter each phase can be.

Considering the Appropriateness of the Changing Criterion Design

The changing criterion design is a useful addition to the behavior analyst's set of tactics for evaluating systematic behavior change. Like the multiple baseline design, the changing criterion design does not require that behavioral improvements be reversed. However, partial reversals to earlier levels of performance enhance the design's capability to demonstrate experimental control. Unlike the multiple baseline design, only one target behavior is required.

Several characteristics of the changing criterion design limit its effective range of applications. The design can be used only with target behaviors already in the subject's repertoire and that lend themselves to stepwise modification. However, this is not as severe a limitation as it might seem. For example, students perform many academic skills to some degree, but not at useful rates. Many of these skills (e.g., solving math problems, reading) are appropriate for analysis with a changing criterion design. Allowing students to progress as efficiently as possible while meeting the design requirements of changing criterion analysis can be especially difficult. As Tawney and Gast (1984) noted, "the challenge of identifying criterion levels that will permit the demonstration of experimental control without impeding optimal learning rates" (p. 298) is problematic with all changing criterion designs.

Although the changing criterion design is sometimes suggested as an experimental tactic for analyzing the effects of shaping programs, it is not appropriate for this purpose. In shaping, a new behavior that initially is not in the person's repertoire is developed by reinforcing responses that meet a gradually changing criterion, called successive approximations, toward the terminal behavior (see Chapter 22). However, the changing response criteria employed in shaping are topographical in nature, requiring different forms of behavior at each new level. The multiple probe design is more appropriate for analyzing a shaping program because each new response criterion (successive approximation) represents a different response class whose frequency of occurrence is not wholly dependent on the frequency of behaviors meeting other criteria in the shaping program. Conversely, the changing criterion design is best suited for evaluating the effects of instructional techniques on stepwise changes in the rate, accuracy, duration, or latency of a single target behavior.

SUMMARY

Multiple Baseline Design

1. In a multiple baseline design, simultaneous baseline measurement is begun on two or more behaviors. After stable baseline responding has been achieved, the independent variable is applied to one of the behaviors while baseline conditions remain in effect for the other behavior(s). After maximum change has been noted in the first behavior, the independent variable is applied in sequential fashion to the other behaviors in the design.
2. Experimental control is demonstrated in a multiple baseline design by each behavior changing when, and only when, the independent variable is applied.
3. The multiple baseline design takes three basic forms: (a) a multiple baseline across behaviors design consisting of two or more different behaviors of the same subject; (b) a multiple baseline across settings design consisting of the same behavior of the same subject in two or more different settings; and (c) a multiple baseline across subjects design consisting of the same behavior of two or more different participants.

Variations of the Multiple Baseline Design

4. The multiple probe design is effective for evaluating the effects of instruction on skill sequences in which it is highly unlikely that the participant's performance on later steps in the sequence can improve without instruction or mastery of the earlier steps in the chain. The multiple probe design is also appropriate for situations in which prolonged baseline measurement may prove reactive, impractical, or too costly.
5. In a multiple probe design, intermittent measures, or probes, are taken on all of the behaviors in the design at the outset of the experiment. Thereafter, probes are taken each time the subject masters one of the behaviors or skills in the sequence. Just prior to instruction on each behavior, a series of baseline measures is taken until stability is achieved.
6. The delayed multiple baseline design provides an analytic tactic in situations in which (a) a planned reversal design is no longer desirable or possible; (b) limited resources preclude a full-scale multiple baseline design; or (c) a new behavior, setting, or subject appropriate for an ongoing multiple baseline analysis becomes available.
7. In a delayed multiple baseline design, baseline measurement of subsequent behaviors is begun sometime after baseline measurement began on earlier behaviors in the design. Only baselines begun while earlier behaviors in the design are still under baseline conditions can be used to verify predictions made for the earlier behaviors.
8. Limitations of the delayed multiple baseline design include the following: (a) The behavior analyst may have to wait too long to modify certain behaviors; (b) baseline phases may contain too few data points, and (c) baselines begun after the independent variable has been applied to earlier behaviors in the design can mask the interdependence (covariation) of behaviors.
9. The nonconcurrent multiple baseline across participants design consists of a series of A-B (baseline-intervention) sequences conducted across subjects at different points in time.
10. The nonconcurrent multiple baseline across participants design is inherently weaker than other multiple baseline design variations. The nonconcurrent design entails two of three elements of baseline logic—prediction and replication—but the absence of concurrent baseline measures does not enable verification of intervention effects.

Assumptions and Guidelines for Using Multiple Baseline Designs

11. Behaviors that are part of multiple baseline designs should be functionally independent (i.e., they do not covary) and share a reasonable likelihood that each will change when the independent variable is applied to it.
12. Behaviors in a multiple baseline design should be measured concurrently and have an equal opportunity of being influenced by the same set of relevant variables.
13. In a multiple baseline design, the independent variable should not be applied to the next behavior until the previous behavior has changed maximally and a sufficient period of time has elapsed to detect any effects on behaviors still in baseline conditions.
14. The length of the baseline phases for the different behaviors making up a multiple baseline design should vary significantly.
15. All other things being equal, the independent variable should be applied first to the behavior showing the most stable level or counter-therapeutic trend of baseline.
16. Conducting a reversal phase in one or more tiers of a multiple baseline design can strengthen the demonstration of a functional relation.

Considering the Appropriateness of Multiple Baseline Designs

17. Advantages of the multiple baseline design include the fact that (a) it does not require withdrawing a seemingly effective treatment; (b) sequential implementation of the independent variable parallels the practice of many teachers and clinicians whose task is to change multiple behaviors in different settings and/or subjects; (c) the concurrent measurement of multiple behaviors allows direct monitoring of generalization of behavior change; and (d) the design is relatively easy to conceptualize and implement.

18. Limitations of the multiple baseline design include the fact that (a) if two or more behaviors in the design covary, the multiple baseline design may not demonstrate a functional relation even though one exists; (b) because verification must be inferred from the lack of change in other behaviors, the multiple baseline design is inherently weaker than the reversal design in showing experimental control between the independent variable and a given behavior; (c) the multiple baseline design is more an evaluation of the independent variable's general effectiveness than an analysis of the behaviors involved in the design; and (d) conducting a multiple baseline design experiment requires considerable time and resources.

Changing Criterion Design

19. The changing criterion design can be used to evaluate the effects of a treatment on the gradual or stepwise improvement of a behavior already in the subject's repertoire.

20. After stable baseline responding has been achieved, the first treatment phase is begun, in which reinforcement is usually contingent on the subject performing at a specified level (criterion). The design entails a series of treatment phases, each requiring an improved level of performance over the previous phase. Experimental control is demonstrated when the subject's behavior closely conforms to the changing criteria.

21. In a range-bound changing criterion design, each intervention sub-phase includes both a lower and an upper criterion within which the subject is expected to perform.

22. Three features combine to determine the potential for a changing criterion design to demonstrate experimental control: (a) the length of phases, (b) the magnitude of criterion changes, and (c) the number of criterion changes. The believability of the changing criterion design is enhanced if a previous criterion is reinstated and the subject's behavior reverses to the level previously observed under that criterion.

Considering the Appropriateness of the Changing Criterion Design

23. Advantages of the changing criterion design are (a) it does not require a withdrawal or reversal of a seemingly effective treatment; and (b) it enables an experimental analysis within the context of a gradually improving behavior, thus complementing the goals of practitioners.

24. Limitations of the changing criterion design are (a) the target behavior must already be in the subject's repertoire, and (b) incorporating the necessary features of the design may impede optimal learning rates.

KEY TERMS

changing criterion design
delayed multiple baseline design
multiple baseline across behaviors design
multiple baseline across settings design
multiple baseline across subjects design
multiple baseline design
multiple probe design
nonconcurrent multiple baseline across participants design
range-bound changing criterion design

MULTIPLE-CHOICE QUESTIONS

1. Simultaneous measurement of two or more behaviors of a single participant describes which type of experimental design?

   a. Multiple baseline across behaviors
   b. Multiple baseline across settings
   c. Multiple baseline across subjects
   d. Multiple baseline across probes

   Hint: (See "Multiple Baseline Design")

2. In a multiple baseline design, the functional relationship between the independent variable and the change in the subject's behavior is determined by:

   a. Systematic introduction and withdrawal across phases
   b. Observation of unchanged behaviors in other tiers
   c. Observation of behavior change in the treated tier
   d. Replication of behavior change with each new criterion

   Hint: (See "Multiple Baseline Design")

3. Which experimental design is most appropriate for evaluating the effect of video modeling on the acquisition of the skills for preparing a meal by adults with developmental disabilities?
   a. Multiple baseline across behaviors
   b. Changing criterion
   c. Multiple probe
   d. Delayed multiple baseline

   Hint: (See "Variations of the Multiple Baseline Design")

4. All of the following are appropriate situations to use a delayed multiple baseline except:
   a. A planned reversal design is no longer possible
   b. The researcher has limited resources to conduct data collection
   c. A reversal design is too challenging to implement
   d. Additional subjects become available after the start of the study

   Hint: (See "Variations of the Multiple Baseline Design")

5. All of the following are advantages of the multiple baseline design except:
   a. Easy to conceptualize
   b. Lack of treatment withdrawal
   c. Mimics common practice of practitioners
   d. Takes less time to implement

   Hint: (See "Assumptions and Guidelines for Using Multiple Baseline Designs")

6. All of the following are limitations of the multiple baseline design except:
   a. It may not reveal the functional relationship, even if one exists
   b. Verification of predicted behavior change must be inferred from other behaviors.
   c. The behavior under study must be within the subject's repertoire already
   d. It is weaker than the reversal design

   Hint: (See "Assumptions and Guidelines for Using Multiple Baseline Designs")

7. When an intervention that is known to be effective is withheld from a subject for achieving a multiple baseline design, this presents a(n) ________ challenge for researchers.
   a. Practical
   b. Ethical
   c. Experimental
   d. Logical

   Hint: (See "Assumptions and Guidelines for Using Multiple Baseline Designs")

8. One method of addressing the practical concern of costly measurement when using a multiple baseline design across settings is to:
   a. Include only one setting in the study
   b. Reduce the number of baseline data points
   c. Use intermittent probes during baseline
   d. Use a changing criterion design instead of the multiple baseline

   Hint: (See "Assumptions and Guidelines for Using Multiple Baseline Designs")

9. When a subject's behavior changes each time a new criterion is introduced, this is an example of
   a. Prediction
   b. Function
   c. Replication
   d. Systematic manipulation

   Hint: (See "Changing Criterion Design")

10. Which of the following is not a design factor to be considered when using a changing criterion design?
    a. Length of phase
    b. Magnitude of behavior change
    c. Number of criterion changes
    d. Verification of behavior change

    Hint: (See "Guidelines for Using the Changing Criterion Design")

ESSAY-TYPE QUESTIONS

1. Explain how a researcher uses prediction and verification of behavior change in a multiple baseline across behaviors design.

   Hint: (See "Multiple Baseline Design")

2. Identify the three types of multiple baseline designs and give an example of when each is appropriate.

   Hint: (See "Multiple Baseline Design")

3. Briefly describe how prediction and replication are demonstrated in the changing criterion design.

   Hint: (See "Operation and Logic of the Changing Criterion Design")

4. Describe the five guidelines for using a multiple baseline design.

   Hint: (See "Assumptions and Guidelines for Using Multiple Baseline Designs")

5. Describe a situation in which the most ethical experimental design would be the multiple probe design.

   Hint: (See "Variations of the Multiple Baseline Design")

NOTES

1. Although most of the graphic displays created or selected for this text as examples of experimental design tactics show data plotted on noncumulative vertical axes, the reader is reminded that repeated measurement data collected within any type of experimental design can be plotted with a variety of graphic displays. For example, Lalli, Zanolli, and Wohn (1994) and Mueller, Moore, Doggett, and Tingstrom (2000) used cumulative graphs to display the data they collected in multiple baseline design experiments; and Raiff, Jarvis, and Dallery (2016) presented the results of a multiple baseline design experiment with a graphic display consisting of a horizontal sequence of boxes in different shades to illustrate participants' performances during baseline and treatment conditions. Students of applied behavior analysis should be careful not to confuse the different techniques for graphically displaying data with tactics for experimental analysis.
2. Functional analysis and automatic reinforcement are described in Chapters 11 and 27, respectively.
3. Time delay, a procedure for transferring stimulus control from a response prompt (e.g., teacher says "apple" while pointing to an apple) to a natural stimulus (e.g., an apple), is described in Chapter 17.
4. In most of these studies, the experimental designs were described as multiple probe designs. However, the graphic displays show that baselines were added in a time-delayed fashion.
5. *Hypothesis*, as we are using the term here, should not be confused with formal hypothesis testing models that use inferential statistics to confirm or reject a hypothesis deduced from a theory. As Johnston and Pennypacker (2009) pointed out, "Researchers do not need to state hypotheses if they are asking a question about nature. When the experimental question simply asks about the relation between independent and dependent variables, there is no scientific reason to make a prediction about what will be learned from the data. Whatever is learned will describe something about that relationship that was presumably not known before. Whether it matches anyone's expectations has nothing to do with its truth or importance (Sidman)." (p. 50)
6. Negative reinforcement and differential reinforcement of other behavior (DRO) are detailed in Chapters 12 and 25, respectively.

CHAPTER 10

Planning and Evaluating Applied Behavior Analysis Research

LEARNING OBJECTIVES

- Explain the importance of the individual subject in behavioral research.
- Explain the importance of flexibility when designing experiments.
- Identify and design systems for evaluating internal validity.
- Identify and design systems for evaluating social validity.
- Identify and design systems for monitoring external validity.
- Explain the importance of evaluating applied behavior analysis research.

Previous chapters outlined considerations and procedures for selecting and defining target behaviors, detailed strategies for designing and implementing measurement systems, presented guidelines for displaying and interpreting behavioral data, and described experimental tactics for revealing whether observed changes in behavior can be attributed to an intervention. This chapter supplements the information described thus far by examining additional questions and considerations that should be addressed when designing, replicating, and evaluating behavioral research. We begin by reviewing the central role of the individual subject in behavioral research, and follow with discussions of the value of flexibility in experimental design, how researchers using single-subject and between-groups designs pursue internal and external validity, and the importance and means of assessing the applied value of treatments and the behavior changes they bring about. The chapter concludes with suggestions for evaluating the believability and usefulness of research in applied behavior analysis.

IMPORTANCE OF THE INDIVIDUAL SUBJECT IN BEHAVIOR ANALYSIS RESEARCH

The research methods of any science must match the defining characteristics of that science's subject matter. Behavior analysis—a science devoted to discovering and understanding the controlling variables of *behavior*—defines its subject matter as the activity of living organisms, a dynamic phenomenon that occurs at the level of the individual organism. It follows that behavior analysis research methods feature direct and repeated measures of the behavior of individual organisms (the only place, by definition, where behavior takes place). This focus on the behavior of individual subjects has enabled applied behavior analysts to discover and refine effective assessment tools and interventions for a wide range of socially significant behavior.

To further explain the importance this focus on the individual subject or client holds for applied behavior analysis, we contrast it with a research model that compares data representing the aggregate measures of different groups of subjects. This between-groups approach to experimental design has predominated "behavioral research" in psychology, education, and other social sciences for decades.

Brief Outline of a Between-groups Experiment

The basic format of a between-groups experiment is as follows.

1. A pool of subjects (e.g., 60 first-grade nonreaders) is selected randomly from the population (e.g., all first-grade nonreaders in a school district) relevant to the research question (e.g., Will the XYZ intensive phonics program improve first-grade nonreaders' ability to decode unpredictable text?).
2. The subjects are divided randomly into two groups: the experimental group and the control group.
3. An initial measure (pretest) of the dependent variable (e.g., score on a test of decoding skills) is obtained for all subjects in the study, the individual pretest scores for the subjects in each group are combined, and the mean and standard deviation are calculated for each group's performance on the pretest.
4. Subjects in the experimental group are then exposed to the independent variable (e.g., 6 weeks of the XYZ program), which is not provided to subjects in the control group.
5. After the treatment program has been completed, a posttest measure of the dependent variable is obtained for all subjects, and the mean and standard deviation posttest scores for each group are computed.[1]
6. The researcher compares any changes in each group's scores from pretest to posttest, applying various statistical tests to the data that enable inferences regarding the likelihood that any differences between the two

groups' performances can be attributed to the independent variable. For example, assuming the mean pretest scores for the experimental and control groups were similar, and the posttest revealed an improved score for the experimental group but not for the control group, statistical analyses would indicate the mathematical probability that the difference was due to chance. When a statistical test limits the role of chance to a predetermined acceptable degree—usually 5% (1 chance in 20) or 1% (1 chance in 100)—the researcher concludes that the independent variable was responsible for changes to the dependent variable (e.g., the experimental group's improvement from pretest to posttest).[2]

Researchers who compare measures of groups of subjects in this way do so for two primary reasons, each of which was introduced in Chapter 7. First, the assumption that averaging the measures of many subjects' performance controls for intersubject variability enables the belief that any changes in performance are brought about by the independent variable. A second rationale for large groups of subjects is the notion that increasing the number of subjects increases the study's external validity. That is, a treatment variable found effective with the subjects in the experimental group will also be effective with other subjects in the population from which the sample subjects were selected.

In the next section we comment on the first reason for the use of groups of subjects—that doing so controls for intersubject variability. We identify four fundamental concerns with typical between-groups designs that bear heavily on experimental reasoning. The assumption of increased generality of findings is discussed later in this chapter in the section on replication.[3]

Group Data May Not Represent the Performance of Individual Subjects

Knowing that the average performance of a group of subjects changed may reveal nothing about the performance of individual subjects. The average performance of subjects in the experimental group may improve, while the performance of some subjects in the group stayed the same, and the performance of others deteriorated. It is even possible for the majority of subjects to show no improvement, for some subjects to get worse, and for a few subjects to improve sufficiently to yield an overall average improvement.

Between-groups studies can show that a treatment is generally effective, that no treatment works with everyone, that people respond differently to the same treatment, and so on. However, the fact that the average performance of a group of subjects improved when they were exposed to a treatment is insufficient reason to adopt it. The factors responsible for one subject's improvement and another's lack of improvement must be discovered. To be most useful, a treatment must be understood at the level at which people contact it and are affected by it: the individual level.

The two graphs in Figure 10.1 illustrate some of the faulty conclusions possible when interpreting a study on the basis of group mean scores. Each graph presents hypothetical data for a two-subject group. The group data showing no change from pretest to posttest in both graphs suggest that the independent variable had no effect on the subjects' behavior. However, the left-hand graph in Figure 10.1 shows that Subject A's performance improved from pretest to posttest while Subject B's behavior deteriorated over the same period.[4] The right-hand graph shows identical pre- and posttest measures for Subjects C and D, while repeated measures between the pretest and posttest reveals significant variability within and between the two subjects.

Group Data Mask Variability

A second problem associated with the mean performance of a group of subjects is that it hides variability in the data. Even if repeated measures of Subject C's and Subject D's behavior between the pre- and posttest had been conducted as shown in Figure 10.1, a researcher who relied on the group's mean performance as the primary indicator of behavior change would be ignorant of the variability that occurred within and between subjects.

When repeated measurement reveals significant levels of variability, an experimental search with the goal of identifying and controlling the factors responsible for the variability is in order. The widespread belief that the effects of uncontrolled variables in a study can be controlled by statistical manipulation of the dependent variable is faulty.

> Statistical control is never a substitute for experimental control. . . . The only way to determine whether or not uncontrolled variables are influencing the data is to inspect the data at the finest available level of decomposition,

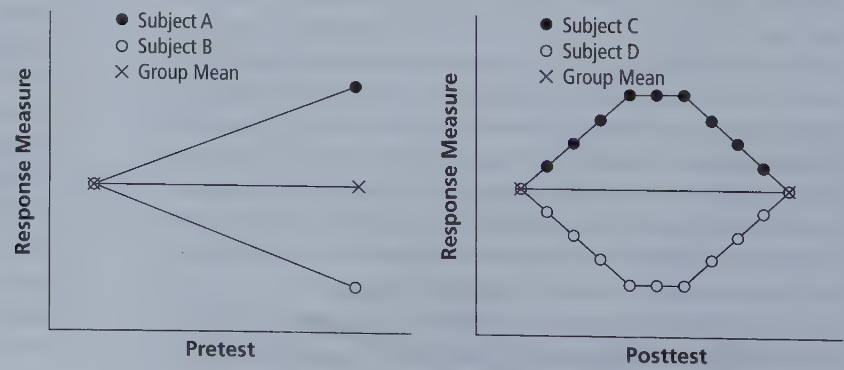


Figure 10.1 Hypothetical data showing that the mean performance of a group of subjects may not represent the behavior of individual subjects.

> usually point-by-point for each individual subject. No purpose is served by combining the data statistically to obscure such effects. (Johnston & Pennypacker, 1980, p. 371)

Attempting to "cancel out" variability through statistical manipulation neither eliminates it nor controls the variables responsible for it. And the researcher who attributes the effects of unknown or uncontrolled variables to chance forgoes the opportunity to identify and analyze important variables. In his monumental work *Tactics of Scientific Research*, Sidman (1960) dealt repeatedly and forcefully with this critical issue.

> To some experimenters, chance is simply a name for the combined effects of uncontrolled variables. If such variables are, in fact, controllable, then chance in this sense is simply an excuse for sloppy experimentation, and no further comment is required. If the uncontrolled variables are actually unknown, then chance is, as Boring (1941) has pointed out, a synonym for ignorance. . . . One of the most discouraging and at the same time challenging aspects of behavioral science is the sensitivity of behavior to a tremendous array of variables. . . . But variables are not canceled statistically. They are simply buried so that their effects cannot be seen. The rationale for statistical immobilization of unwanted variables is based on the assumed random nature of such variables. . . . Not only is the assumption of randomness with respect to the uncontrolled variables an untested one but it is also highly improbable. There are few, if any, random phenomena in the behavioral world. (pp. 45, 162–163)

Sidman (1960) also commented on an experimenter's use of statistics in an attempt to deal with troublesome sequence effects.

> He has a neat trick up his sleeve. By averaging together the data for both subjects under Condition A, and again under Condition B, he "cancels out" the order effect, and completely bypasses the problem of irreversibility. By a simple arithmetical operation, two subjects have become one, and a variable has been eliminated.
>
> It has not, in fact, gone anywhere. Numbers may be made to disappear by adding and subtracting them from each other. Five apples minus three apples are two apples. The numbers are easily changed by a few strokes of the pen, but some eating has to be done before the apples themselves will vanish. (p. 250)

The "eating" that must be done to control the effects of any variable can be accomplished in only two ways: (a) holding the variable constant throughout the experiment, or (b) isolating the suspected factor as an independent variable and manipulating its presence, absence, and/or value during the experiment.

Group Data Do Not Represent Real Behavioral Processes

Skinner (1938) contended that researchers must demonstrate behavior–environment relations at the level of the individual organism or risk discovering synthetic phenomena that represent mathematical, not behavioral, processes. Sidman (1960) described the challenges faced by an investigator interested in the relation between the number of times a behavior is reinforced during initial acquisition and responding during a subsequent extinction condition. A single-case approach would entail presenting an individual subject with a series of acquisition-extinction events, each acquisition condition consisting of a different number of reinforcements. The subject's performance during each successive extinction phase would be confounded by the subject's experience with all previous reinforcement-then-extinction events. Behavior during extinction would decrease to zero more rapidly as a function of that cumulative experience, rather than the number of reinforcements obtained in the acquisition phase immediately prior to each extinction test.

A between-groups design solution to the problem that no individual subject can be exposed to differing amounts of reinforcement during an initial acquisition phase would be to assign multiple subjects to different groups, expose subjects in each group to different values of the independent variable (e.g., 5, 10, and 20 reinforcements), and compare the mean rate of responding by subjects in each group during a subsequent extinction condition. Such a design may reveal a relation between number of reinforcements during initial acquisition and responding during extinction, but . . .

> The function does not represent a behavioral process. The use of separate groups destroys the continuity of cause and effect that characterizes an irreversible behavioral process. . . . If it proves impossible to obtain an uncontaminated relation between the number of reinforcements and resistance to extinction in a single subject, because of the fact that successive extinctions interact with each other, then the "pure" relation simply does not exist. The solution is . . . to direct our research toward the study of behavior as it actually exists . . . [and] not be deceived into concluding that the group type of experiment in any way provides a more adequately controlled or more generalizable substitute for individual data. (Sidman, 1960, p. 53)

Acknowledging that studies comparing group data describe "some kind of order in the universe" and may form the basis of a science, Sidman (1960) noted they:

> cannot, however, be a science of individual behavior except of the crudest sort. . . . It is a science of averaged behavior of individuals who are linked together only by the averaging process itself. Where this science fits in the scheme of natural phenomena is a matter of conjecture. My own feeling is that it belongs to the actuarial statistician, and not the investigator of behavior processes. (pp. 274–275)

Between-groups Designs Lack Intrasubject Replication

A fourth weakness of the between-groups research model is that the power of replicating effects within and across individual subjects is lost. One of the great strengths of single-case experimental designs is the convincing demonstration of functional relations made possible by replication of treatment effects within the designs. Although multiple subjects typically participate in

single-case studies and investigators often describe and display the data for all subjects as a group, data from individual subjects are the basis for interpreting experimental effects. Applied behavior analysts are wise to heed Johnston and Pennypacker's (1980) admonition: "An effect that emerges *only* after individual data have been combined is probably artifactual and not representative of any real behavioral processes" (p. 257).

This discussion should not be interpreted to mean that the overall performance of groups of subjects cannot, or should not, be studied with the research strategies and tactics of applied behavior analysis. Improving overall performance of a group is socially significant in many applied situations (Donaldson, Fisher, & Kahng, 2017; Friman, Finney, Rapoff, & Christophersen, 1985; Luyben, 2009). Brothers, Krantz, and McClannahan (1994) used a multiple baseline across three settings to evaluate an intervention that increased the number of pounds of office paper recycled by 25 school staff members. Cope, Allred, and Morsell (1991) employed a multiple treatment reversal design to reveal which of three experimental conditions most effectively reduced illegal parking by drivers of 558 vehicles in spaces designated for individuals with disabilities at a supermarket.

Still, it is important to remember that group data may not represent the performance of individual participants, and vice versa. Lloyd, Eberhardt, and Drake (1996) compared the effects of group versus individual reinforcement contingencies within the context of collaborative group study conditions on quiz scores by students in a Spanish language class. The results showed that the group contingencies resulted in higher mean quiz scores for the class as a whole compared to the individual contingencies condition. However, overall benefits at the class level were mitigated by differential results for individual students. When group results do not represent individual performances, researchers should supplement group data with individual results, ideally in the form of graphic displays (e.g., Lloyd et al., 1996; Ryan & Hemmes, 2005).

In some instances, the behavior analyst may not be able to control potential subjects' access to the experimental setting and contingencies or even be able to identify who the subjects are. The dependent variable must then consist of responses emitted by individuals who enter the experimental setting. This approach is used frequently in community-based behavior analysis research. For example, group data have been collected and analyzed on such dependent variables as recycling on a university campus (O'Conner, Lerman, & Fritz, 2010), car pooling by university students (Jacobs, Fairbanks, Poche, & Bailey, 1982), motorists yielding at pedestrian crosswalks (Crowley-Koch, Van Houten, & Lim, 2011), the use of child safety belts in shopping carts (Barker, Bailey, & Lee, 2004), and reducing graffiti on restroom walls (Watson, 1996).

IMPORTANCE OF FLEXIBILITY IN EXPERIMENTAL DESIGN

On one level, an effective experimental design is any sequence of independent variable manipulations that produces data that are interesting and convincing to the researcher and the audience. In this context *design* is appropriate as a verb as well as a noun; the effective behavioral researcher must actively *design* each experiment so that each achieves its own unique *design*. The prototype designs presented in Chapters 8 and 9 are examples of analytic tactics that afford a form of experimental reasoning and control that has proven effective in advancing the scientific understanding of a wide range of phenomena of interest to applied behavior analysts.

An investigator should not fall into the trap of believing that a certain type of research question, treatment, behavior, or setting indicates which analytic tactic(s) should be employed. While the experimental designs in published studies exploring similar research questions should be mined for methodological ideas, no ready-made experimental designs await selection, nor is there a set of rules that must be followed. Sidman (1960) was adamant in his warning regarding the undesirable effects of researchers' believing in the existence of a given set of rules for experimental design.

> The examples may be accepted as constituting a set of rules that must be followed in the design of experiments. I cannot emphasize too strongly that this would be disastrous. I could make the trite statement that every rule has its exception, but this is not strong enough. Nor is the more relaxed statement that the rules of experimental design are flexible, to be employed only where appropriate. The fact is that *there are no rules of experimental design*. (p. 214)

We agree with Sidman. The student of applied behavior analysis should not be led to believe that any of the analytic tactics described in Chapters 8 and 9 constitute experimental designs per se.[5] Still, we believe that it is useful to present the most commonly used analytic tactics in design form for two reasons. First, the vast majority of studies that have advanced the knowledge base of applied behavior analysis have used experimental designs featuring one or more of these analytic tactics. Second, we believe examining specific examples of isolated experimental tactics and their application is an important step in learning the assumptions and strategic principles that guide the selection and arrangement of analytic tactics into experimental designs that effectively and convincingly address the research questions at hand.

Experimental Designs Combining Analytic Tactics

An experimental design that combines analytic tactics may allow a more convincing demonstration of experimental control than a design using a single analytic tactic. For example, withdrawing the treatment variable (a return to baseline) and then reapplying it within one or more tiers in a multiple baseline design may enable the researcher to determine the existence of a functional relation between the independent variable and each behavior, setting, or subject of the multiple baseline, in addition to analyzing the effectiveness of the independent variable across the tiers (e.g., Barker et al., 2004; Donaldson et al., 2017; Heward & Eachus, 1979; Miller & Kelley, 1994 [see Figure 28.14]; Zhou, Goff, & Iwata, 2000).

Whereas a design with a single analytic tactic may function like a belt that effectively holds up a pair of jeans, a design

that combines analytic tactics is akin to adding a pair of suspenders that more firmly answers the research or adds a bit of stylistic flair. However, beginning researchers should not assume that designs with multiple tactics are inherently better than single-tactic designs. A design featuring multiple analytic tactics may be weaker than a single-tactic design, if the additional analytic tactics overshadow the powerful answer provided by data with a simple, yet elegant single-tactic design.

Experimental designs combining multiple baseline, reversal, and/or multielement tactics can provide the basis for comparing the effects of two or more independent variables. Haring and Kennedy (1990) used multiple baseline across settings and reversal tactics in an experiment comparing the effectiveness of time-out and differential reinforcement of other behavior (DRO) on the occurrence of problem behaviors by two secondary students with severe disabilities (see Figure 10.2).[6] Sandra and Raff each frequently engaged in repetitive, stereotypic problem behaviors (e.g., body rocking, loud vocalizations, hand flapping, spitting) that interfered with classroom and community activities. In addition to assessing the effects of the time-out and DRO interventions against a no-treatment baseline condition, the design also enabled the researchers to conduct two comparisons of the relative effects of each treatment during an instructional task and leisure context. The design enabled Haring and Kennedy to discover that the time-out and DRO interventions produced different outcomes depending on the activity context in which they were applied. For both students, DRO was more effective than time-out in suppressing problem behavior in the task context; the opposite results were obtained in the leisure context, where time-out suppressed problem behavior and DRO proved ineffective.

Experimenters have also incorporated alternating treatments into experimental designs containing multiple baseline elements. For example, Ahearn, Kerwin, Eicher, Shantz, and

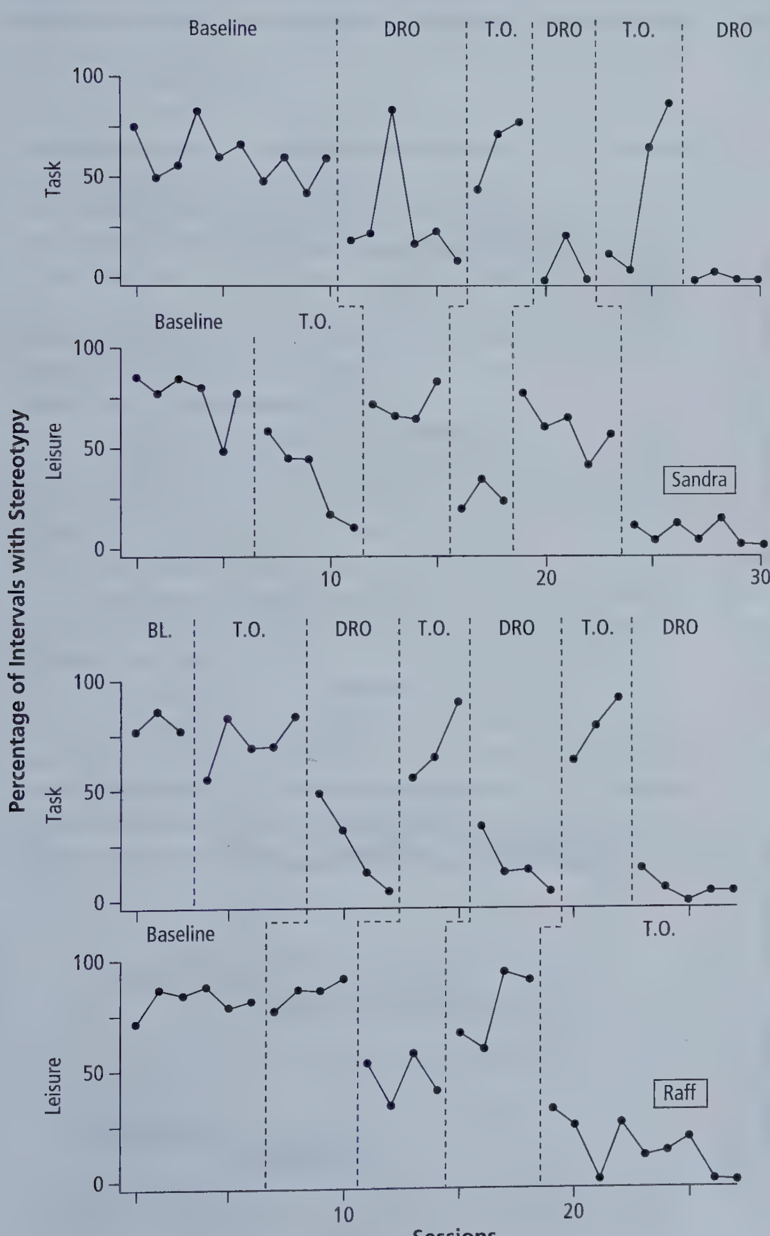


Figure 10.2 Experimental design employing multiple baselines across settings and reversal tactics counterbalanced across two subjects to analyze the effects of time-out (TO) and differential reinforcement of other behavior (DRO) treatment conditions.

Based on "Contextual Control of Problem Behavior" by T. G. Haring and C. H. Kennedy, 1990, *Journal of Applied Behavior Analysis*, *23*, pp. 239–240. Copyright 1990 by the Society for the Experimental Analysis of Behavior, Inc.

Swearingin (1996; see Figure 12.5) evaluated the relative effects of two treatments for food refusal in an alternating treatments design implemented in a multiple baseline across subjects format. Likewise, Mcgee Krantz, and McClannahan (1985) evaluated the effects of several procedures for teaching language to children with autism with an experimental design that incorporated alternating treatments within a multiple baseline across behaviors component that was, in turn, nested within an overall multiple baseline across subjects format. Zanolli and Daggett (1998) investigated the effects of reinforcement rate on the spontaneous social initiations of socially withdrawn preschoolers with an experimental design consisting of multiple baseline, alternating treatments, and reversal tactics.

Sisson and Barrett (1984) incorporated multiple probe, multielement, and a multiple baseline across behaviors tactics in a design comparing the effects of two language-training procedures (see Figure 10.3). The design enabled the investigators to discover the superiority of the total communication method for these two children, as well as the fact that direct application of the treatment was required for learning to occur on specific sentences. Results for a third subject revealed a functional relation of the same form and direction as that found for the two children whose results are shown in Figure 10.3, but one not so strongly in favor of the total communication procedure.

Component Analyses

When a behavioral intervention consists of multiple components, it is termed a **treatment package**. An effective toilet training program, for example, might entail replacing diapers with underwear or training pants, a schedule of sitting on the toilet for various durations, increased fluid intake, reinforcement for proper eliminations, and contingencies for accidents (Greer, Neibert, & Dozier, 2016 [see Figure 6.7]). Multicomponent interventions are common for treating complex feeding disorders. Treatment packages for food packing (holding or pocketing food in the oral cavity without swallowing) have included combinations of the following components: presenting highly preferred food on the spoon behind a bite of previously packed food; removing packed food and placing it on the child's tongue; a liquid chaser; inserting food

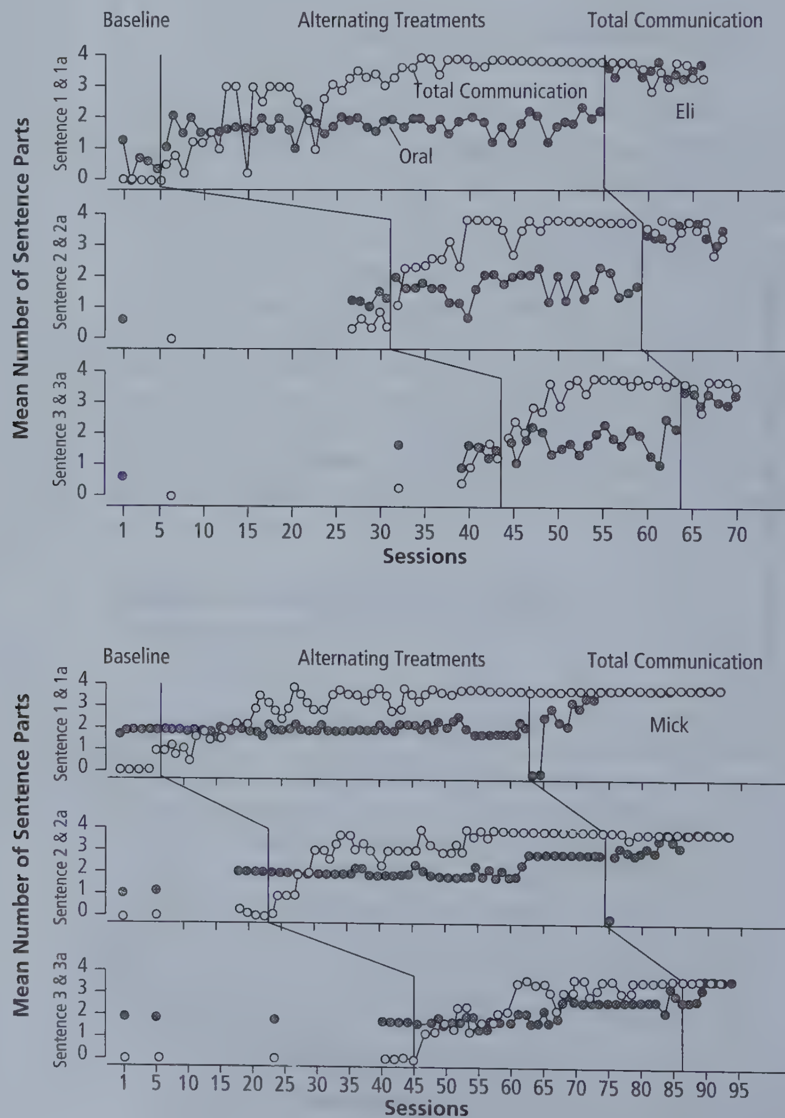


Figure 10.3 Experimental design employing a multiple probe, multielement, and multiple baseline analytic tactics.

"Alternating Treatments Comparison of Oral and Total Communication Training with Minimally Verbal Retarded Children", L. A. Sisson and R. P. Barrett, 1984, *Journal of Applied Behavior Analysis*, *17*, p. 562. Reproduced with permission of John Wiley & Sons Inc.

into the child's mouth with a spoon, then flipping the spoon 180° and dragging the bowl of the spoon toward the lips to deposit the food on the child's tongue; removing packed food with a spoon; facilitating swallowing with a light touch on the tongue's posterior; response cost for packing or expulsion; and various forms of positive reinforcement (e.g., preferred toy) and negative reinforcement (e.g., break from feeding) for swallows (Levin, Volkert, & Piazza, 2014; Sharp, Odom, & Jaquess, 2012).

Practitioners and researchers create, implement, and evaluate treatment packages for many reasons. Here are several: A practitioner may believe two or more research-based interventions will be more effective than any of the component interventions in isolation. Interventions consisting of multiple components may be effective across a wider range of settings or participant characteristics. An intervention known to be mildly effective in isolation may, when implemented in conjunction with other research-based interventions, yield an additive, value-enhancing effect. In an effort to change a behavior that has resisted numerous interventions, a practitioner might combine some or all of the failed interventions, perhaps adding new ones, into a treatment package.

When desired behavior change coincides with or follows implementation of a treatment package, an analysis of whether or not a functional relation exists between the package and behavioral improvement is in order. Other important questions can then be pursued: Must every component of a treatment package be implemented to achieve clinically meaningful behavior change? If not, which components are necessary? Is one component, or a subset of components, sufficient to bring about behavior change? How does the presence or absence of each component impact overall effectiveness?

A **component analysis** is any experiment designed to identify the active elements of a treatment package, the relative contributions of different components in a treatment package, and/or the necessity and sufficiency of treatment components. The effectiveness of a treatment package requires the presence of all *necessary* components. A *sufficient* component, or subset of components, ensures treatment effectiveness in the absence of any other components.

In their review of component analyses using single-subject experimental designs, Ward-Horner and Sturmey (2010) described the potential outcomes of component analyses using reversal designs of a treatment package consisting of two components (Y and Z). In their hypothetical examples shown in Figure 10.4, Component Y is praise and Component Z is a primary reinforcer.

> *Outcome 1* suggests that Y is necessary and sufficient, because Y is effective whether or not Z is present.
> *Outcome 2* suggests an *additive effect* because the sum of the components produced behavior changes greater than either component alone.
> *Outcome 3* illustrates a *multiplicative effect* because neither Y nor Z is effective in isolation, but the combination of these components produces substantial behavior change.
> *Outcome 4* indicates components Y and Z are sufficient but neither component is necessary, whereas Outcomes 2 and 3 indicate that neither Component Y nor Component Z is sufficient but both are necessary.
> *Outcome 5* is challenging to interpretations: (a) the possibility that presenting the components together establishes relations between or among themselves that influence subsequent effects when presented alone; (b) perhaps Y is effective only because of its prior presentation with Z; or (c) Y may be effective only when preceded by Z, but Y may be ineffective when it precedes Z. (Adapted from Ward-Horner & Sturmey, 2010, p. 687)

Ward-Horner and Sturmey (2010) described two methods for conducting component analyses. In a **drop-out component analysis**, the investigator presents the treatment package and then systematically removes components. If the treatment's effectiveness wanes when a component is removed, the researcher has identified a necessary component.

> The main advantage of this approach is that the target behavior will improve immediately, within the first or second experimental phase, and that the subsequent removal of components may provide information on the components necessary to maintain treatment goals (L. J. Cooper et al., 1995). One disadvantage is that the behavioral effects of combining components might mask the effectiveness of individual components, which makes it difficult to determine the necessity and sufficiency of components. (p. 687)

A drop-out analysis by Wacker and colleagues (1990) revealed that both components of a training package consisting of functional communication training and consequences for inappropriate behavior were necessary for three individuals with severe disabilities and behavior problems.

An **add-in component analysis** assesses components individually or in combination before the complete treatment package is presented. The add-in method can identify sufficient components. Ward-Horner and Sturmey (2010) noted that sequence and floor or ceiling effects may mask the effects of components added in toward the end of the analysis.

After data from a reversal design demonstrated the effectiveness of a treatment package of noncontingent reinforcement, response cost, and response blocking for wearing prescription glasses by three individuals with intellectual disabilities, DeLeon and colleagues (2008) used add-in or drop-out and component analyses to determine which components were sufficient and/or necessary for each participant.

Our intent in describing several experiments that combined analytic tactics or analyzed the relative effects of treatment components is not to offer any of these studies as model designs. They are presented instead as illustrations of the infinite number of experimental designs that are possible by arranging different combinations and sequences of independent variable manipulations. In every instance the most effective (i.e., convincing) experimental designs are those that use an ongoing evaluation of data from individual subjects as the basis for employing the three elements of baseline logic—prediction, verification, and replication.

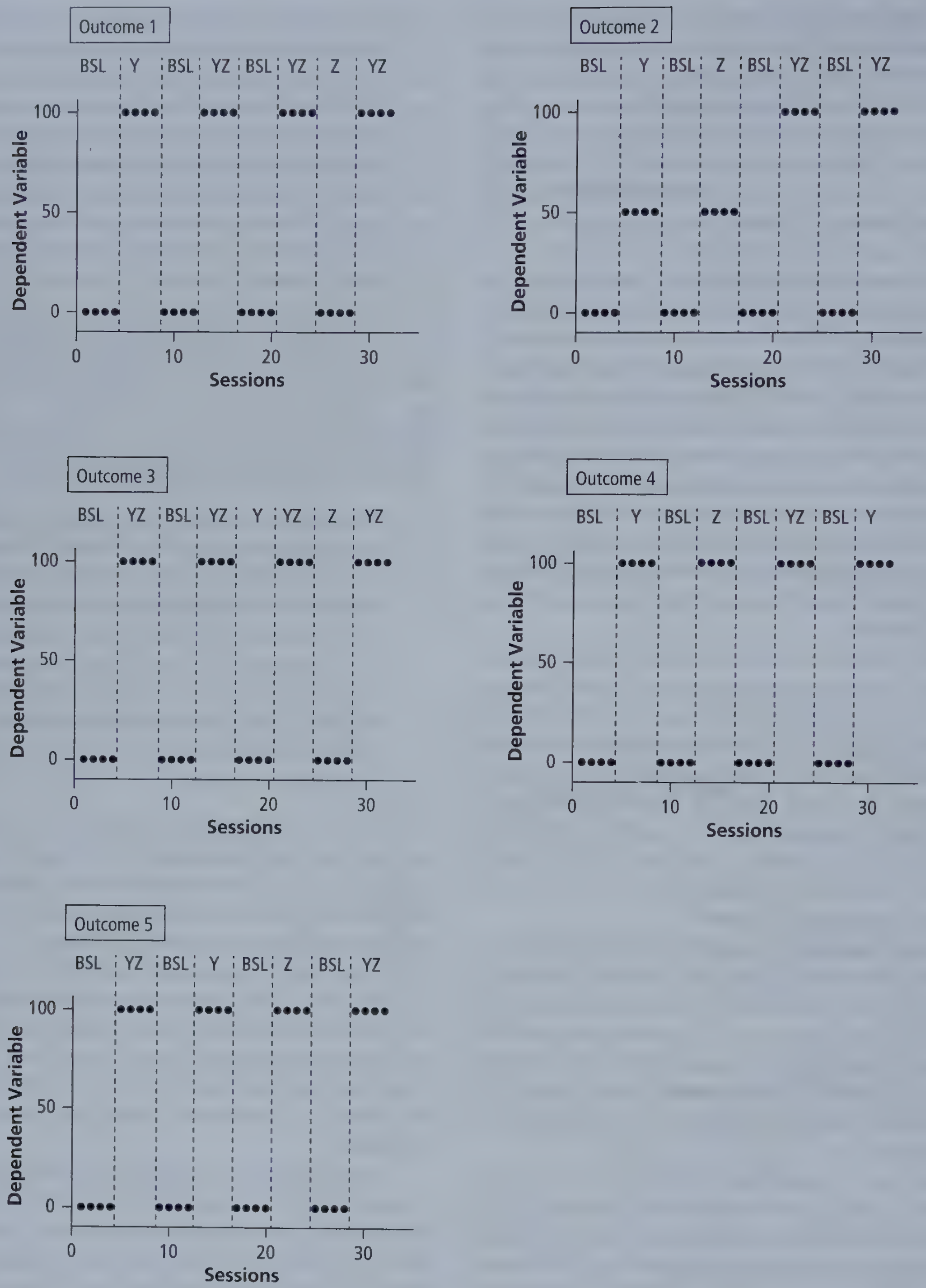


Figure 10.4 Potential outcomes of component analyses using reversal designs for a two-component treatment package.

From "Component Analyses Using Single-Subject Experimental Designs: A Review," by J. Ward-Horner and P. Sturmey, 2010, *Journal of Applied Behavior Analysis*, *43*, p. 690. Copyright 2010 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

INTERNAL VALIDITY: CONTROLLING POTENTIAL SOURCES OF CONFOUNDING IN EXPERIMENTAL DESIGN

An experiment is interesting and convincing, and yields the most useful information for application, when it provides an unambiguous demonstration that the independent variable was solely responsible for the observed behavior change. Experiments that demonstrate a clear functional relation have a high degree of internal validity. An experimental design's strength is determined by the extent to which it (a) demonstrates a reliable effect (i.e., repeated manipulation of the independent variable produces a consistent pattern of behavior) and (b) eliminates or reduces the likelihood that factors other than the independent variable produced the behavior change (i.e., controls for confounding variables).

Implicit in the term *experimental control*, which is often used to signify a researcher's ability to reliably produce a specified behavior change by manipulating an independent variable, is the idea that the researcher controls the subject's behavior. However, "control of *behavior*" is inaccurate because the experimenter can control only some aspect of the subject's *environment*. Therefore, the level of experimental control obtained by a researcher refers to the extent to which she controls all relevant variables in a given experiment. The researcher exerts this control within the context of an experimental design that, even though carefully planned at the outset, takes its ultimate form from the researcher's ongoing examination and response to the data.

As noted in Chapter 7, an effective experimental design simultaneously reveals a reliable functional relation between independent and dependent variables (if one exists) and minimizes the likelihood that the observed behavior changes are the result of unknown or uncontrolled variables. An experiment has high internal validity when changes in the dependent variable are demonstrated to be a function only of the independent variable. When planning an experiment and later when examining the data from an ongoing study, the investigator must always be alert to threats to internal validity. Uncontrolled factors known or suspected to have exerted influence on the dependent variable are called *confounding variables* (see Chapter 7). Much of a researcher's efforts during the course of a study are aimed at identifying and controlling confounding variables.

The attainment of steady state responding is the primary means by which applied behavior analysts assess the degree of experimental control. Separating the effects of the independent variable from the effects of a potentially confounding variable requires clear, empirical evidence that the potentially confounding variable is no longer present, has been held constant across experimental conditions, or has been isolated for manipulation as an independent variable. Any experiment can be affected by a nearly unlimited number of potential confounds; and as with every aspect of experimental design, there are no set rules for identifying and controlling confounding variables to which researchers can turn. However, common and likely sources of confounding can be identified as well as tactics that can be considered to control them. Confounding variables can be viewed as related primarily to one of four elements of an experiment: subject, setting, measurement of the dependent variable, and independent variable.

Subject Confounds

A variety of subject variables can confound the results of a study. *Maturation*, which refers to changes that take place in a subject over the course of an experiment, is a potential confounding variable. For example, a subject's improved performance during the later phases of a study may result from physical growth or the acquisition of academic, social, or other behaviors and may be unrelated to manipulations of the independent variable. Experimental designs that incorporate rapidly changing conditions or multiple introductions and withdrawals of the independent variable over time usually control for maturation effectively.

In most applied behavior analysis research, a subject is in the experimental setting and contacting the contingencies implemented by the investigator for only a portion of the day. As in any study, the assumption is made that a subject's behavior during each session will be primarily a function of the experimental conditions in effect. In reality, however, a subject's behavior may also be influenced by events that occur outside the experiment. For example, suppose that the rate of contributions to a class discussion is the dependent variable. Now suppose that a student who has been contributing to discussions at a high rate gets into a fight in the lunchroom just prior to a session and then emits substantially fewer contributions compared to his level of responding in previous sessions. This change in the student's behavior may, or may not, be a result of the lunchroom fight. If the lunchroom fight coincided with a change in the independent variable, it would be especially difficult to detect or separate any effects of the experimental conditions from those of the extra-experimental event.

Although the researcher may be aware of some events that are likely causes of variability during a study, many other potential confounds go undetected. Repeated measurement is both the control for and the means to detect the presence and effects of such variables. The uncontrolled variables responsible for a subject's having a "bad day" or an unusually "good day" are particularly troublesome in research designs with few and/or widely spaced measurements of the dependent variable. This is one of the major weaknesses of using data from pretest-posttest comparisons to evaluate the effects of a treatment program.

Because between-groups experiments are predicated on subjects' similarity in relevant characteristics (e.g., gender, age, ethnicity, cultural and linguistic background, current skills and background knowledge), they are vulnerable to confounding by differences among subjects. Concern that the characteristics of one or more subjects may confound an experiment's results is generally not an issue in the single-case experiments. First, a person should participate in a study because she will benefit if the target behavior is changed successfully. Second, a subject's idiosyncratic characteristics cannot confound a study using a true intrasubject experimental design. With the exception of the multiple baseline across subjects analysis design, each participant in a single-case study serves as her own control, a strategy that guarantees identically matched subjects in all experimental conditions *because those subjects are the same person*. Third, the external validity of the findings of a single-subject analysis is not dependent on the extent to which the subject(s) shares certain characteristics with others. The extent to which a functional relation applies to other subjects is established by replicating the experiment with different subjects (replication is discussed later in this chapter).

Setting Confounds

Most applied behavior analysis studies are conducted in natural settings where a host of variables are beyond the investigators' control. Studies in natural settings are more prone to confounding by uncontrolled events than are studies conducted in laboratories where extraneous variables can be eliminated

or held consistent. Even so, the applied experimenter is not without resources to mitigate the detrimental effects of setting confounds. For instance, when the applied researcher observes that an uncontrolled event has coincided with changes in the data, he should hold all possible aspects of the experiment constant until repeated measurement again reveals stable responding. If the unplanned event appears to have had a robust effect on the target behavior, or is otherwise of interest to the investigator, and is amenable to experimental manipulation, the investigator should treat the unplanned event as an independent variable and explore its possible effects experimentally.

Applied researchers concerned about setting confounds must also be on the lookout for "bootleg" reinforcement within and outside the experimental situation. A good example of how a setting confound operates occurs when, unbeknownst to the experimenter, subjects have access to the same items or events to be used as putative reinforcers in the study. In such a case, the potential effectiveness of those consequences as reinforcers may be diminished considerably.

Measurement Confounds

Chapters 4 and 5 discussed factors that should be considered in designing an accurate and nonreactive measurement system. Still, numerous sources of confounding may exist within a well-planned measurement system. For instance, data might be confounded by observer drift, the influence of the experimenter's behavior on observers, and/or observer bias. Although admittedly difficult to accomplish in applied settings where observers often see the independent variable being implemented, keeping observers naive to the conditions and expected outcomes of an experiment reduces the potential of confounding by observer bias. On a related note, when observers score permanent products, the products should not contain identifying marks that indicate who produced each product and under what experimental conditions it was produced. Having observers score papers from baseline and treatment conditions in randomized order reduces the likelihood that observer drift or bias will confound the data within one treatment condition phase. (This procedure is more suitable to controlling for drift or bias by observers conducting postexperiment accuracy or interobserver agreement [IOA] assessments.)

Unless a completely unobtrusive measurement system is devised (e.g., a covert system using one-way mirrors, or observations conducted at some distance from the subject), reactivity to the measurement procedure must always be considered as a possible confound. To offset this possible confound, the experimenter must maintain baseline conditions long enough for any reactive effects to run their course and stable responding to be obtained. If reactivity to measurement produces undesirable effects (e.g., aggressive behavior, cessation of productivity) and a more unobtrusive measurement procedure cannot be devised, intermittent probes should be considered. Practice, adaptation, and warm-up effects can also confound measures, especially during the initial stages of baseline. Again, the proper procedure is to continue baseline conditions until stable responding is obtained or variability is reduced to minimal levels. Intermittent probes should not be used for baseline measurement of behaviors for which practice effects would be expected. This is because, if the target behavior is susceptible to practice effects, those effects will occur during the intervention condition when more frequent measures are conducted, thereby confounding any effects of the independent variable.

Independent Variable Confounds

Most independent variables are multifaceted; that is, there is usually more to a treatment condition than the specific variable of interest to the investigator. For example, the effects of a token economy on students' academic productivity may be confounded by variables such as the personal relationship between the students and the teacher who delivers the tokens, social interactions associated with delivering and exchanging tokens, the expectation of teacher and students that performance will improve when the token system is implemented, and so on. If the intent is to analyze the effects of token reinforcement per se, these potentially confounding variables must be controlled.

Schwarz and Hawkins (1970) provided a good example of a control procedure for ruling out an aspect associated with a treatment as responsible for behavior change. The researchers evaluated the effects of token reinforcement on three maladaptive behaviors of an elementary student who was described as severely withdrawn. During treatment, the therapist and the girl met each day after school and viewed a videotape that had been made earlier that day of the student's classroom behavior. The therapist also administered tokens contingent on the girl's videotaped behavior displaying progressively fewer occurrences of maladaptive behaviors.

Schwarz and Hawkins recognized an independent variable confound potentially lurking in the design of their study. They reasoned that if improvement occurred as a function of the treatment, the question would remain whether the student's behavior had improved because the therapist-delivered positive attention and rewards improved her self-concept, which in turn changed her maladaptive behaviors in the classroom, which were symptomatic of her poor self-concept. In that case, Schwarz and Hawkins could not be certain that the contingent tokens played an important role in changing the behavior. Anticipating this possible confound, Schwarz and Hawkins controlled for it in a simple and direct way. Following baseline, they implemented a condition in which the therapist met with the girl each day after school and provided her with social attention and token reinforcement contingent on improvements in handwriting. During this control phase, the three target behaviors—face touching, slouching, and low voice volume—showed no change, thereby increasing their confidence in a conclusion that the girl's ultimate behavioral improvements during the subsequent intervention phases were due to the token reinforcement.

When medical researchers design experiments to test the effects of a drug, they use a technique called a **placebo control** to separate effects that may be produced by a subject's perceived expectations of improvement because of taking the drug apart from the effects actually produced by the drug. In the typical between-groups design, the subjects in the experimental group receive the real drug, and subjects in the control group receive

placebo pills. Placebo pills contain an inert substance, but they look, feel, and taste exactly like the pills containing the real drug being tested.

Applied behavior analysts have also employed placebo controls in single-subject experiments. For example, in their study evaluating a pharmacological treatment of the impulsivity by students with attention-deficit hyperactivity disorder (ADHD), Neef, Bicard, Endo, Coury, and Aman (2015) had a pharmacist prepare placebos and medications in identical gelatin capsules in 1-week supplies for each child. Neither the students nor the observers knew if a child had taken the medication or had taken the placebos. When neither the subject(s) nor the observers know whether the independent variable is present or absent from session to session, this type of control procedure is called a **double-blind control**. A double-blind control procedure eliminates confounding by subject expectations, parent and teacher expectations, differential treatment by others, and observer bias.

Treatment Integrity/Procedural Fidelity

Researchers must ensure that the independent variable is implemented exactly as planned and that no other unplanned variables are administered inadvertently along with the planned treatment. **Treatment integrity** refers to the extent to which the independent variable is implemented as planned. Although often used interchangeably with treatment integrity, **procedural fidelity** more accurately refers to the extent to which procedures in all conditions of an experiment, including baseline, are correctly implemented (Ledford & Gast, 2014).

Low treatment integrity invites a major source of confounding into an experiment, making it difficult, if not impossible, to interpret the results with confidence. Data from an experiment in which the independent variable was administered improperly, applied inconsistently, conducted piecemeal, and/or delivered in overdose or underdose form often lead to conclusions that—depending on the results obtained—represent either a false positive (claiming a functional relation when no such relation exists) or a false negative (failing to detect a functional relation when one actually does exist). If a functional relation is apparent from the analysis of the data, one cannot be sure whether the treatment variable as described by the experimenter was responsible or whether the effects were a function of extraneous, uncontrolled elements of the intervention as it was actually applied. However, it may be equally erroneous to interpret the failure to produce significant behavior change as evidence that the independent variable is ineffective. In other words, had the independent variable been implemented as planned, it might have been effective.

Numerous threats to treatment integrity exist in applied settings (Billingsley, White, & Munson, 1980; Ledford & Gast, 2014; Peterson, Homer, & Wonderlich, 1982). Experimenter bias can influence the researcher to administer the independent variable in such a way that it enjoys an unfair advantage over baseline or comparative conditions. **Treatment drift** occurs when the application of the independent variable differs from the way it was applied at the study's outset. Treatment drift can occur when the complexity of the independent variable makes it difficult for practitioners to implement all elements consistently over the course of an experiment. Contingencies influencing the behavior of those responsible for implementing the independent variable can also produce treatment drift. For example, a practitioner might implement only those aspects of a procedure she favors and implement the full intervention only when the investigator is present.

The following discussion and examples of methods for ensuring and measuring treatment integrity also apply to ensuring consistency of procedures during baseline and nontreatment comparison conditions (i.e., procedural fidelity).

Precise Operational Definition. Achieving a high level of treatment integrity begins with developing a complete and precise operational definition of the treatment procedures. Besides providing the basis for training the persons who will implement an intervention and judging the level of treatment integrity attained, operational definitions of treatment conditions are requisite for meeting the technological dimension of applied behavior analysis (Baer, Wolf, & Risley, 1968). An investigator's failure to provide explicit operational definitions of the treatment variable hampers the dissemination and proper use of the intervention by practitioners, and makes it difficult for other researchers to replicate and ultimately validate the findings.

Gresham, Gansle, and Noell (1993) recommended that descriptions of the independent variable be judged by the same standards of explicitness used in assessing the quality of definitions of the dependent variable. That is, they should be clear, concise, unambiguous, and objective. More specifically, Gresham and colleagues also suggested that treatments be operationally defined in each of four dimensions: verbal, physical, spatial, and temporal. They used Mace, Gresham, Ivancic, and O'Brien's (1986) definition of a time-out procedure as an example of an operational definition of an independent variable.

> (a) Immediately following the occurrence of a target behavior (temporal dimension), (b) the therapist said "No, go to time-out" (verbal dimension), (c) led the child by the arm to a prepositioned time-out chair (physical dimension), and (d) seated the child facing the corner (spatial dimension). (e) If the child's buttocks were raised from the time-out chair or if the child's head was turned more than 45° (spatial dimension), the therapist used the least amount of force necessary to guide compliance with the time-out procedure (physical dimension). (f) At the end of 2 min (temporal dimension), the therapist turned the time-out chair 45° from the corner (physical and spatial dimensions) and walked away (physical dimension). (pp. 261–262)

Simplify, Standardize, and Automate. Simplifying and standardizing the independent variable and providing criterion-based training and practice for the people who will be responsible for implementing it enhance treatment integrity. Treatments that are simple, precise, and brief, and that require relatively little effort, are more likely to be delivered with consistency than those that are not. Simple, easy-to-implement techniques also have a higher probability of being accepted and used by practitioners than those that are not, and thus possess a certain

degree of self-evident social validity. Simplicity is, of course, a relative concern, not a mandate; effecting change in some socially important behaviors may require the application of intense, complex interventions over a long period and may involve many people. Baer (1987) made this point succinctly when he stated:

> Long problems are simply those in which the task analysis requires a series of many behavior changes, perhaps in many people, and although each of them is relatively easy and quick, the series of them requires not so much effort as time, and so it is not arduous but merely tedious. (pp. 336–337)

Neither researchers nor practitioners need be thwarted or dismayed by complex interventions; they do, however, need to realize the treatment integrity implications. All things being equal, a simple and brief treatment will probably be applied more accurately and consistently than will a complex and extended one.

To ensure the consistent implementation of the independent variable, experimenters should standardize as many of its aspects as cost and practicality allow. They can accomplish the standardization of treatment in a variety of ways. When a treatment requires a complex and/or extended sequence of behaviors, a script for the person administering it may improve the accuracy and consistency with which the independent variable is applied. For example, Heron, Heward, Cooke, and Hill (1983) used a scripted lesson plan to ensure that a classwide peer tutor training program was implemented consistently across groups of children.

If automating the intervention will not compromise it in any way, researchers might consider "canning" the independent variable so that an automated device can be used for its delivery. Although a tutor training video in Heron and colleagues' (1983) study would have eliminated any potential confounding caused by the teacher's slightly different presentations of the lesson from group to group and across sets of tutoring skills, using a canned presentation would also have eliminated desired interactive and personal aspects of the training program. Some treatment variables are well suited to automated presentation in that automation neither limits the desirability of the treatment nor seriously reduces its social validity in terms of acceptability or practicability (e.g., use of video programs to model residential energy conservation).

Training, Practice, and Feedback. Training and practice in properly implementing the independent variable provides the person(s) who will be responsible for conducting the treatment or experimental sessions with the necessary skills and knowledge to carry out the treatment. It would be a mistake for the researcher to assume that a person's general competence and experience in the experimental setting (e.g., a classroom) guarantee correct and consistent application of an independent variable in that setting (e.g., implementing a peer-mediated tutoring program).

As stated earlier, scripts detailing treatment procedures and cue cards or other devices that remind and prompt people through steps of an intervention can be helpful. Researchers should not, however, assume that providing a procedural manual or script to intervention agents will ensure a high degree of treatment integrity. Mueller and colleagues (2003) found that a combination of verbal instructions, modeling, and/or rehearsal was required for parents to implement pediatric feeding protocols with a high level of treatment integrity.

Howard and DiGennaro Reed (2014) found that after written instructions produced little or no improvement in trainers' implementation of a dog obedience training procedure with shelter volunteers, a single viewing of a video that included correct and incorrect examples of the routine "produced marked improvement" in the correct implementation of the training procedure. Nevertheless, a brief one-on-one training session consisting of modeling, descriptive praise, and feedback until 100% integrity was obtained in a role-play was required for trainers to reach satisfactory levels of treatment integrity (means of 88%, 97%, and 93% for three trainers over 6 to 14 sessions) (Figure 10.5).

Rating the integrity with which video models implement a treatment can also improve the fidelity with which the raters implement the treatment. After an initial 30-min training session in which trainer-reviewed written protocol of functional analysis (FA) procedures produced moderate levels of treatment fidelity (below 85% accuracy) by 17 of 20 graduate students, Field, Frieder, McGee, Peterson, and Duinkerken (2015) had the students rate the fidelity with which video models implemented the FA procedures. The treatment fidelity of 16 of the 17 students who participated in the observation and rating of video models increased further. In a final condition, 7 participants whose execution of FA procedures was still below mastery observed and rated a 2-min video of themselves conducting the intervention in a previous session. The self-observation and rating resulted in 5 of the 7 students achieving further increases in treatment integrity.

Various prompting tactics might promote treatment integrity. In a parametric analysis of teacher praise rates on the on-task behavior of elementary students with autism, the teacher wore a MotivAider®, which vibrated at preset intervals of 1, 2, and 8 per minute to prompt praise statements (Kranak, Alber-Morgan, & Sawyer, 2017).

Craig (2010) reported an interesting method for assessing the treatment integrity of a self-monitoring and self-administering intervention on reduced nail biting. Tokens were exchanged for evening leisure time contingent on no occurrences of nail biting during daily, 1-hour intervals. Craig described the procedure as follows:

> I recorded the time of each occurrence of nail biting and the time of each token reinforcer delivery. Using a cellular camera phone, I took digital photographs (with timestamps) of the condition of my fingernails at the beginning of each DRO interval, after each token delivery, and after each occurrence of nail biting (see Figure 1). During token exchange periods each evening, my cohabitant compared the times of the occurrences of nail biting and of token deliveries to the timestamps of each photograph. If there was no noticeable difference in the condition of my fingernails between the photograph from the beginning of an interval and the photograph taken upon token delivery, token delivery was recorded as "correct." Tokens

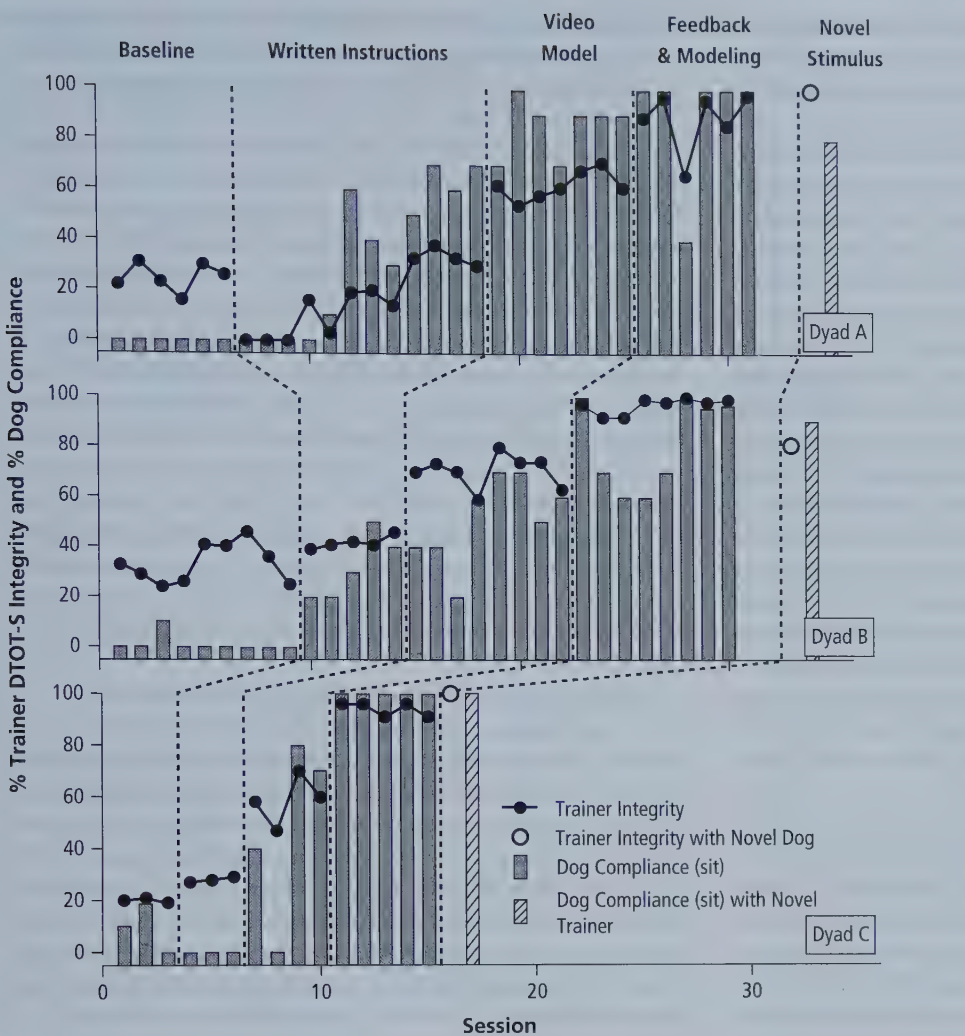


Figure 1. Trainer integrity of discrete-trial obedience training for sit (DTOT-s) and dog compliance with sit instruction.

Figure 10.5 Graphic display illustrating treatment integrity data and dog compliance.

Based on "Training shelter volunteers to teach dog compliance" by V. J. Howard and F. D. DiGennaro Reed, 2014, *Journal of Applied Behavior Analysis, 47*, p. 351. Copyright 2014 by the Society for the Experimental Analysis of Behavior, Inc.

> associated with incorrect deliveries would have been removed; incorrect token delivery did not, however, occur during the course of this study. (p. 39)

Performance feedback and self-monitoring have also been shown to improve the integrity with which practitioners and parents implement behavior support plans and explicit teaching techniques (e.g., Codding, Feinberg, Dunn, & Pace, 2005; Pinkelman & Horner, 2016; Plavnick, Ferreri, & Maupin, 2010; Sarokoff & Strumey, 2004; Sanetti, Luiselli, & Handler, 2007).

Assessing and Reporting Treatment Integrity. Simplification, standardization, training, and practice help increase the degree of treatment integrity, but do not guarantee it. When any reasonable doubt exists about the correct and consistent application of the independent variable, investigators should provide data on the accuracy and reliability of the independent variable. Procedural fidelity data reveal the extent to which the actual implementation of all experimental conditions over the course of a study matches their descriptions in the method section of a research report.[7]

Even though effective control of the independent variable is requisite to an internally valid experiment, applied behavior analysts have not always made sufficient efforts to ensure treatment integrity. Two reviews of articles published in the *Journal of Applied Behavior Analysis (JABA)* from 1968 to 1990 found that the majority of authors did not report data assessing the degree to which the independent variable was properly and consistently applied (Gresham et al., 1993; Peterson et al., 1982). Peterson and colleagues noted that a "curious double standard" had developed in applied behavior analysis in which data on the interobserver agreement of dependent variable measures were required for publication, but such data were seldom provided or required for the independent variable.

Although more recent reviews of treatment integrity data in articles published in *JABA* and other behavioral journals have found some improvement, the overall results are disappointing (e.g., Andzik, Cannella-Malone, & Sigafoos, 2016; Ledford & Wolery, 2013; Wheeler, Baggett, Fox, & Blevins, 2006). For example, McIntyre, Gresham, DiGennaro, and Reed (2007) reported that only 30% of 152 studies of school-based interventions published in *JABA* from 1991 to 2005 provided treatment integrity data. These authors judged 69 (45%) of those studies "to be at high risk for treatment inaccuracies in

that information on the implementation of treatments or the assessment of independent variables was not included but should have been" (p. 664) per guidelines for monitoring or measuring treatment (Peterson et al., 1982).

Treatment integrity data have value beyond their critical role in helping reviewers, practitioners, and consumers judge the believability of a study's results. Increasingly, data documenting that the independent variable was implemented as described is criterion for inclusion in systematic reviews of literature conducted by professional and government organizations. Treatment integrity is one of five dimensions by which the National Autism Center (NAC; 2015) judges the scientific merit of a study (the others are research design, measurement of the dependent variable, eligibility and description of participants, and measure of generalization of treatment effects). For a study to achieve the highest score on the NAC's Scientific Merit Rating Scale, treatment integrity must be measured during a minimum of 25% of sessions, at or above 80% accuracy, with IOA of ≥80%. An expert panel tasked by the U.S. Department of Education to establish parameters for considering fidelity of intervention implementation in single-case design studies noted that "Because interventions in single-case studies are applied over time, continuous measurement of implementation is a relevant consideration" (What Works Clearinghouse, 2018, p. A-6).

Methods for assessing and increasing the accuracy and believability of measures of the dependent variable (see Chapter 5) are fully applicable to the collection of procedural fidelity data. Importantly, observation and recording of the independent variable provide the experimenter with data indicating whether calibration of the treatment agent is necessary (i.e., bringing the intervention agent's behavior into agreement with the true value of the independent variable). Observation and calibration give the researcher an ongoing ability to use retraining and practice to ensure a high level of treatment integrity over the course of an experiment.

Figure 10.6 shows the data collection form trained observers used to collect treatment integrity data in a study evaluating the effects of different qualities and durations of reinforcement for problem behavior, compliance, and communication within a treatment package for escape-maintained problem behavior (Van Norman, 2005). The observers viewed videotapes of randomly selected sessions representing approximately one third to one half of all sessions in each condition and phase of the study. The percentage of treatment integrity for each condition was calculated by dividing the number of steps the experimenter completed correctly during a session by the total number of steps completed.

Graphic displays of treatment integrity data may help researchers and consumers of a study judge the intervention's effectiveness (e.g., Kranak et al., 2017). Detailed information on assessing, promoting, and reporting procedural fidelity can be found in a special issue of the *Journal of Behavioral Education* (DiGennaro Reed & Codding, 2014) and Ledford and Gast (2014).

This overview of sources of potential confounding variables is, of necessity, incomplete. A complete inventory of all possible threats to the internal validity of experimental research would be well beyond the scope of this text. And presenting such a list might suggest that a researcher need only control for the variables listed and not worry about anything else. In truth, the list of potential confounds is unique to every experiment. The effective researcher is the one who questions and probes the influence of as many relevant variables as possible. No experimental design can control all potential confounds; the challenge is to reduce, eliminate, or identify the influence of as many potentially confounding variables as possible.

Video Clip # 1-1AL   Rater Initials: E. B.   Date: 7/6/05

Phase 1/A (SD)

Procedural Steps	Opportunities	Correct	% Correct	Yes	No	N/A
1. The instructor delivers a task prompt at the beginning of the session, e.g., "it's time to work" or similar.				Y	N	N/A
2. If the participant does not respond, the instructor represents the choice by replacing materials or restating the contingencies.	—	—	—	Y	N	N/A
3. The break card (or similar) is presented simultaneously (within 3 seconds) with work task materials.	𝍸 𝍸 𝍸 |	𝍸 𝍸 𝍸 |	16/16	Y	N	N/A
4. Following work choice (touching materials associated with work) a. Removes the task materials b. Presents a timer with green colored cue card c. Provides access to high preference items d. Engages in play with the participant for 1 minute	𝍸 ||	𝍸 ||	7/7	Y	N	N/A
5. Following break requests a. Removes the task materials b. Presents timer with yellow cue card c. Provides access to moderately preferred tangible items and neutral commenting for 30 seconds	𝍸 |||	𝍸 |||	8/8	Y	N	N/A
6. Following problem behavior within 10 s the instructor a. Removes the task/play materials b. Presents a timer with red cue card c. Provides no attention or tangible items for 10 seconds				Y	N	N/A
7. Which side was the break card (or similar) presented (*the participant's* R = right or L = left) for each choice presentation.	R R L L L R L R R L R L R L L R					

Figure 10.6 Example of a form used to record treatment integrity data.

Adapted from *The Effects of Functional Communication Training, Choice Making, and an Adjusting Work Schedule on Problem Behavior Maintained by Negative Reinforcement*, by R. K. Van Norman, 2005, p. 204. Unpublished doctoral dissertation. Columbus, OH: The Ohio State University. Used by permission.

SOCIAL VALIDITY: ASSESSING THE APPLIED VALUE OF BEHAVIOR CHANGES AND THE TREATMENTS THAT ACCOMPLISH THEM

In his landmark article "Social Validity: The Case for Subjective Measurement or How Applied Behavior Analysis Is Finding Its Heart," Montrose Wolf (1978) proposed the then "radical concept that clients (including parents and guardians of dependent people, and even those whose taxes support social programs) must understand and admire the goals, outcomes, and methods of an intervention" (Risley, 2005, p. 284). Wolf recommended that the **social validity** of an applied behavior analysis study be assessed in three ways: the social significance of the behavior change goals, the appropriateness of the intervention, and the social importance of the results.

Validating the Social Importance of Behavior Change Goals

The social validity of behavior change goals begins with a clear description of those goals.

> To assess the social importance of goals, the researcher must be precise about the goals of the behavior change effort at the levels of (a) the broad social goal (e.g., improved parenting, enhanced social skills, improved cardiovascular health, increased independence), (b) the categories of behavior hypothesized to be related to the broad goal (e.g., parenting—providing instructional feedback, using time-out, etc.), and/or (c) the responses that comprise the behavioral category of interest (e.g., using time-out—directing the child to a location away from other people, instructing the child to "sit out" for a specified duration, etc.). Social validation may be conducted for any of these levels of goals. (Fawcett, 1991, pp. 235–236)

Experts can inform the selection of socially valid target behaviors. Kelley and Miltenberger (2016), for example, identified the task analyses of correct rider positions for horseback-riding skills "by consulting with an expert instructor as well as comparing many photos of expert riders considered to have good position" (p. 141). In their studies on the effects of behavioral teaching strategies on performance by high school football players, Harrison and Pyles (2013) and Stokes, Luiselli, and Reed (2010) targeted tackling skills recommended by the American Football Coaches Association (1995). In another study teaching football skills to high school athletes, Stokes, Luiselli, Reed, and Fleming (2010) validated their 10-step task analysis for pass-blocking by consulting with experienced offensive linemen coaches (see Figure 21.5).

Persons who use the prospective skill in natural environments can help researchers identify socially valid target behaviors. In their study teaching sequences of basketball skills to an adolescent with disabilities, Lambert and colleagues (2016) asked "groups of adults who we observed playing games of 'pick up' basketball . . . to rate the utility of each sequence during a typical game of 'pick up' basketball" (p. 206).

Van Houten (1979) suggested two basic approaches for determining socially valid goals: (a) assess the performance of persons considered competent and (b) experimentally manipulate different levels of performance to determine empirically which produces optimal results. The performance of typical performers can be used to identify and validate behavior change goals and target levels of performance. To arrive at a socially valid performance criterion for a social skills training program for two adults with disabilities who worked in a restaurant, Grossi, Kimball, and Heward (1994) observed four restaurant employees without disabilities over a period of 2 weeks to determine the extent to which they acknowledged verbal initiations from coworkers. Results from these observations revealed that the employees without disabilities acknowledged an average of 90% of initiations directed toward them. This level of performance was selected as the goal for the two target employees in the study.

Warren, Rogers-Warren, and Baer (1976) provide a good example of testing the effects of different levels of performance to determine socially valid outcomes. The researchers assessed the effect of different frequencies of children's offers to share play materials with their peers on the peers' reactions to those offers. They found that peers accepted offers to share most consistently when those offers were made at a middle frequency—that is, not too frequently, not too seldom.

Validating the Social Acceptability of Interventions

Several scales and questionnaires for obtaining consumers' opinions of the acceptability of behavioral interventions have been developed. The *Intervention Rating Profile* (IRP-15) is a 15-item Likert-type scale for assessing teachers' perceptions of the acceptability of classroom interventions (Martens, Witt, Elliott, & Darveaux, 1985). The *Behavior Intervention Rating Scale*, a revision and extension of the IRP-15, includes nine additional items on treatment effectiveness (Elliot & Brock Treuting, 1991). The *Treatment Acceptability Rating Form* (TARF) consists of 20 questions with which parents rate the acceptability of behavioral treatments (Reimers & Wacker, 1988). For a review of these and six other treatment acceptability instruments, see Carter (2007).

Figure 10.7 shows the experimenter-modified version of the TARF Van Norman (2005) used to obtain treatment acceptability information from each participant's parents, teachers, therapists, and behavior support staff. Although some of the people whose opinions were being sought had witnessed or watched a video of the intervention being used with the student, the following description of the intervention was read to each consumer before he or she was asked to answer each of the questions:

> First we conducted an assessment to find out what motivated Zachary to engage in challenging behavior(s) such as throwing materials, hitting people, and dropping to the floor. We found that Zachary engaged in challenging behavior(s), at least in part, in order to escape or avoid task demands.
>
> Next, we taught Zachary to ask for a break as a replacement behavior for challenging behavior by using physical prompting and attaching the response of asking for

Figure 10.7 Examples of questions adapted from the Treatment Acceptability Rating Form—Revised (Reimers and Wacker, 1988) to obtain consumers' opinions of the acceptability of intervention procedures used to treat challenging behaviors of secondary students with severe disabilities.

Treatment Acceptability Rating Form—Revised (TARF-R)

1. How clear is your understanding of the suggested procedures?

______	______	______	______	______	______	______
Not at all clear			Neutral			Very clear

2. How acceptable do you find the strategies to be regarding your concerns about the identified learner?

______	______	______	______	______	______	______
Not at all acceptable			Neutral			Very acceptable

3. How willing are you to implement the suggested procedures as you heard them described?

______	______	______	______	______	______	______
Not at all willing			Neutral			Very willing

4. Given the learner's behavior issues, how reasonable do you find the suggested procedures?

______	______	______	______	______	______	______
Not at all reasonable			Neutral			Very reasonable

5. How costly will it be to implement these strategies?

______	______	______	______	______	______	______
Not at all costly			Neutral			Very costly

11. How disruptive will it be to your classroom to implement the suggested procedures?

______	______	______	______	______	______	______
Not at all disruptive			Neutral			Very disruptive

13. How affordable are these procedures?

______	______	______	______	______	______	______
Not at all affordable			Neutral			Very affordable

14. How much do you like the proposed procedures?

______	______	______	______	______	______	______
Do not like them at all			Neutral			Like them very much

17. How much *discomfort* is your learner likely to experience as a result of these procedures?

______	______	______	______	______	______	______
No discomfort at all			Neutral			Very much discomfort

19. How willing would you be to change your classroom routine to implement these procedures?

______	______	______	______	______	______	______
Not at all willing			Neutral			Very willing

20. How well will carrying out these procedures fit into your classroom routine?

______	______	______	______	______	______	______
Not at all well			Neutral			Very well

From *The Effects of Functional Communication Training, Choice Making, and an Adjusting Work Schedule on Problem Behavior Maintained by Negative Reinforcement*, by R. K. Van Norman, 2005, pp. 248–256. Unpublished doctoral dissertation. Columbus, OH: The Ohio State University. Used by permission.

a break to access to a highly preferred item, lots of attention, and a long duration break (3 min).

Then we gave Zachary the choice to ask for work by simply touching the work materials (essentially engaging in the first step of the task) and getting access to highly preferred items, attention, and long duration break (1 min) or asking for a break and getting access to moderately preferred items for a shorter duration break (30 sec). At any time during this procedure if Zachary engaged in problem behavior he was given a 10 s break with no attention and no activities/items.

Finally, we continued to give Zachary the choice to ask for work, a break or engage in problem behavior, however now we required Zachary to comply with a greater number of task-related instructions before he was given access to the highly preferred activities, attention, and a 1 min break. Each session we increased the number of task-related instructions that were given and needed to be complied with before access to the highly preferred break.

Physical prompting was only used during the initial phase to teach Zachary new responses, specifically how to ask for a break and how to ask for work. Otherwise Zachary was making all independent choices as they were presented. (p. 247)

Some investigators present graphic displays of treatment acceptability data. Kurti and Dallery (2013) administered a treatment acceptability questionnaire to 12 adults who participated in Internet-based contingency management intervention to increase walking. Participants ranked 10 items about the intervention (e.g., easy to use, fun, I would recommend to others) using a visual analogue scale ranging from 0 (not at all) to 100 (extremely). The researchers used a graph to present the treatment acceptability results, showing the mean, range, and individual participant ratings for each item (see Figure 10.8).

Preference for interventions can be assessed directly by letting practitioners or clients choose which of multiple treatments to implement or receive. After training four parents and one teacher to implement each of three reinforcement-based interventions for problem behavior to 90% treatment integrity, Gabor, Fritz, Roath, Rothe, and Gourley (2016) provided a sequence of choice sessions in which each caregiver selected which intervention to implement. Jowett Hirst, Dozier, and Payne (2016) assessed preschool children's preference for token economy procedures based on differential reinforcement or response cost with a "choice" phase in which only the treatment selected by each participant was implemented (see Figure 8.17 and accompanying discussion).

Hanley and colleagues have used concurrent chains experimental designs (see Chapter 8) to assess clients' treatment preferences (e.g., Hanley, Piazza, Fisher, Contrucci, & Maglieri, 1997; Luczynski, & Hanley, 2009; Heal, Hanley, & Layer, 2009). After reversal and multielement designs revealed that functional

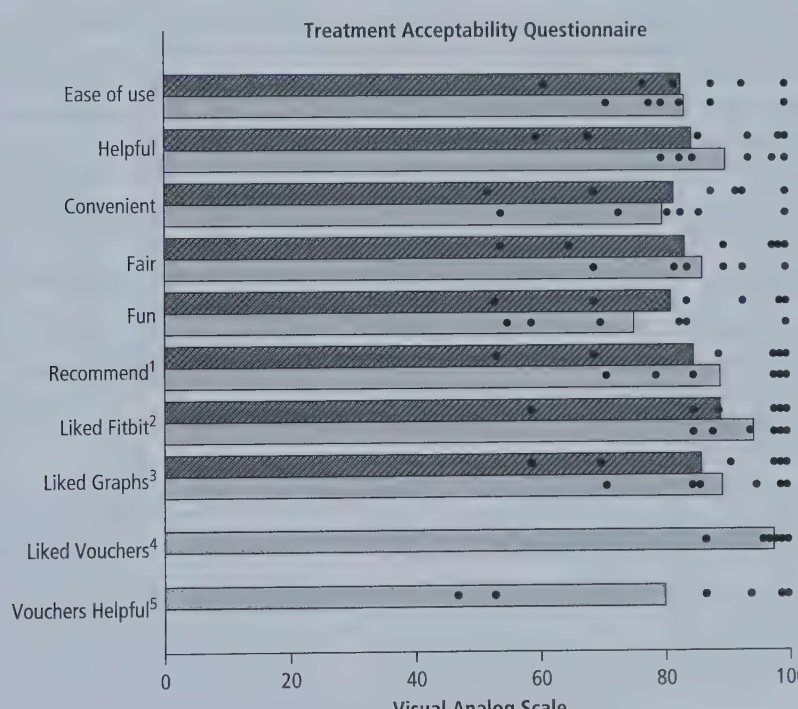


Figure 2. Experiment 1 (gray bars) and Experiment 2 (dark gray bars) participants' responses to the treatment acceptability questionnaire, where bars represent the mean rating for each item (1 = *I would recommend the Internet-based walking intervention to others*, 2 = *I liked using the Fitbit to monitor my progress*, 3 = *I liked seeing my progress on the graph*, 4 = *I liked earning vouchers*, 5 = *Earning vouchers helped me increase my activity levels*).

Figure 10.8 Example of graphic reporting treatment acceptability data.

Based on "Internet-based Contingency Management Increases Walking in Sedentary Adults" by A. N. Kurti and J. Dallery, 2013, *Journal of Applied Behavior Analysis*, *46*, p. 574. Copyright 2013 by the Society for the Experimental Analysis of Behavior, Inc.

communication training (FCT) with punishment was the most effective of three functional-based treatments in reducing two children's severe problem behavior, the children's choices during the concurrent chains condition showed they each preferred functional communication training (FCT) with punishment over FCT alone, or punishment only (Hanley, Piazza, Fisher, & Maglieri, 2005; see Figure 8.18 and accompanying discussion).

The social acceptability of a proposed intervention can be judged by the extent it meets the standards of best practice (Peters & Heron, 1993) and the ethical, legal, and professional standards of relevant learned and professional societies. In general, these standards require the use of evidence-based, and least restrictive practices, full disclosure to all participants, informed consent, and the right to discontinue treatments that the participants believe are not beneficial or therapeutic.

Validating the Social Importance of Behavior Changes

Applied behavior analysts assess the social validity of outcomes with a variety of methods: (a) have consumers rate the social validity of participants' performance, (b) ask experts to evaluate participants' performance, (c) compare participants' performance to that of a normative sample, (d) assess with standardized assessment instruments, and (e) test participants' performance in the natural environment.

Consumer Opinion

The most frequently used method for assessing social validity is to ask consumers, including subjects or clients whenever possible, whether they thought behavior changes occurred during the study or program, and if so, whether they thought those behavior changes were important and valuable. Figure 10.9 shows the questionnaire Van Norman (2005) used to obtain opinions of consumers (i.e., the subjects' parents, teachers, and instructional aides; school administrators; behavior support staff; an occupational therapist; a school psychologist; and a psychology aide) on the social validity of results of an intervention designed to reduce challenging behavior maintained by escape. Van Norman created a series of 5-minute video clips from randomly selected before-intervention and after-intervention sessions and placed the clips in random order on a CD. The social validity evaluators did not know whether each clip

Figure 10.9 Form for obtaining consumer opinions of the social validity of results of an intervention used to treat challenging behavior of secondary students with severe disabilities.

Social Validity of Results Questionnaire

Directions: Video Clip # _____

Please view the video and then circle one of the five choices that best describes the extent to which you agree or disagree with each of the three statements below.

1. The student is engaged in academic or vocational work tasks, sitting appropriately (bottom in seat), and attending to the teacher or materials.

1	2	3	4	5
Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

2. The student is engaged in challenging behavior and not attending to the teacher or materials.

1	2	3	4	5
Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

3. The student appears to have a positive affect (e.g., smiling, laughing).

1	2	3	4	5
Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

Comments about the student's behavior in the video clip: ______________________

____

____

General comments about this student's behavior: ______________________

____

____

Name (optional): ______________________

Relation to the learner in video (optional): ______________________

Based on *"The Effects of Functional Communication Training, Choice Making, and an Adjusting Work Schedule on Problem Behavior Maintained by Negative Reinforcement"* by R. K. Van Norman, 2005, p. 252. Unpublished doctoral dissertation. Columbus, OH: The Ohio State University.

represented a before- or after-intervention session. After viewing each clip, the consumer completed the questionnaire shown in Figure 10.9.

Expert Evaluation

Experts can be called on to judge the social validity of some behavior changes. For example, as one measure of the social validity of changes in the unaided notetaking skills of high school students with learning disabilities in social studies lectures as a result of having been exposed to teacher-prepared guided notes, White (1991) asked 16 secondary social studies teachers to rate the students' baseline and postintervention lecture notes on three dimensions: (1) accuracy and completeness compared to lecture content; (2) usefulness for study for tests over the lecture content; and (3) how the notes compared to those taken by typical general education students. (The teachers did not know whether each set of notes they were rating was from baseline or the postintervention condition.)

Howard and DiGennaro Reed (2014) recruited 186 certified dog trainers from across the United States to evaluate the social validity of behavior changes in dog shelter volunteers and the dogs they were paired with during the study. The researchers compiled brief baseline (before training) and intervention (after training) videos of one of the volunteer+dog dyads. The experts watched the videos online, and using a 9-item SurveyMonkey questionnaire, the experts rated the appropriateness and effectiveness of the dog obedience training procedures used by the shelter volunteer before and after training and whether the before- or after training dog would be more adoptable, better in a home with children, and a better dog for first-time pet owners.

To assess the social validity of changes in participants' horseback-riding skills, Kelley and Miltenberger (2016) showed two expert riders and instructors three videos from the end of baseline and three videos from the end of the intervention for each participant. The experts rated each video with a Likert-type scale (1 = strongly disagree, 9 = strongly agree) according to whether the rider was in the correct position based on the riding style target for the rider shown in the video. The experts who viewed the videos in random order were blind to whether each video was from baseline or intervention.

Normative Comparison

Van den Pol and colleagues (1981) used the performance of a sample of typical fast-food restaurant customers to assess the social validity of the posttraining performance of the young adults with disabilities whom they had taught to order and pay for a meal without assistance. The researchers observed 10 randomly selected, typical customers who ordered and ate a meal in fast-food restaurants and recorded the accuracy with which these customers performed each step of a 22-step task analysis. The students' performance at the follow-up probe equaled or exceeded that of the customers in the normative sample in all but 4 of 22 specific skills.

Using normative samples to assess the social validity of behavior change is not limited to posttreatment comparisons. Comparing subjects' behavior to ongoing probes of the behavior of a normative sample provides a formative assessment measure of how much improvement has been made and how much is still needed. Grossi and Heward (1998) used the typical rate at which employees without disabilities performed restaurant tasks (e.g., scrubbing pots and pans, loading a dishwasher, mopping and sweeping floors) as a socially valid performance standard for restaurant trainees with developmental disabilities (see Figure 29.3).

In their study evaluating behavioral coaching strategies on the offensive line pass-blocking skills of high school football athletes, Stokes, Luiselli, Reed, and Fleming (2010) recorded the three highest-rated starting offensive linemen from videotaped games from the previous season. The mean correct performance on the 10-step task analysis by the three lineman was 80% (range, 70% to 90%). The researchers "adopted a range of 70% to 90% as our acceptable performance criterion" (p. 466).

Rhode, Morgan, and Young (1983) provide an excellent example of formative social validity assessment. In their study, token reinforcement and self-evaluation procedures were used to improve the classroom behavior of six students with behavior disorders. The overall goal of the study was to help the students improve their appropriate classroom behavior (e.g., following classroom rules, completing teacher-assigned tasks, volunteering relevant responses) and decrease inappropriate behavior (e.g., talking out, noncompliance, aggression) so that they would be accepted and successful in regular (general education) classrooms. At least once per day throughout the course of the 17-week study, the researchers randomly selected classmates in the regular classrooms for observation. The same observation codes and procedures that were used to measure the six target students' behavior were used to obtain the normative sample data.

Figure 10.10 shows the mean and range of the six students' appropriate behavior during each condition and phase of the study compared to the normative sample. (Individual graphs showing the percentage of appropriate behavior in the resource and regular classroom of all six subjects in each of nearly 90 sessions were also included in Rhode and colleagues' article.) During baseline, the six boys' levels of appropriate behavior were well below those of their peers without disabilities. During Phase I of the study, in which the subjects learned to self-evaluate, their behavior in the resource room improved to a level matching that of their regular classroom peers. However, when the subjects were in the regular classroom during Phase I, their behavior compared poorly with that of the other students in the normative sample. As Phase II progressed, which involved various strategies for generalization and maintenance of the treatment gains, the mean level of appropriate behavior by the six students matched that of their peers without disabilities, and variability among the six students decreased (except for one subject who exhibited no appropriate behavior on one session in the next-to-last condition).

Standardized Tests

Although standardized tests provide indirect measures of the behaviors typically targeted by applied behavior analysts, policymakers and the public view scores from such tests as important indicators of performance. Improvements in intelligence quotient (IQ) test scores by children who participated in the pioneering studies by Lovaas and colleagues (Lovaas, 1987; McEachin,

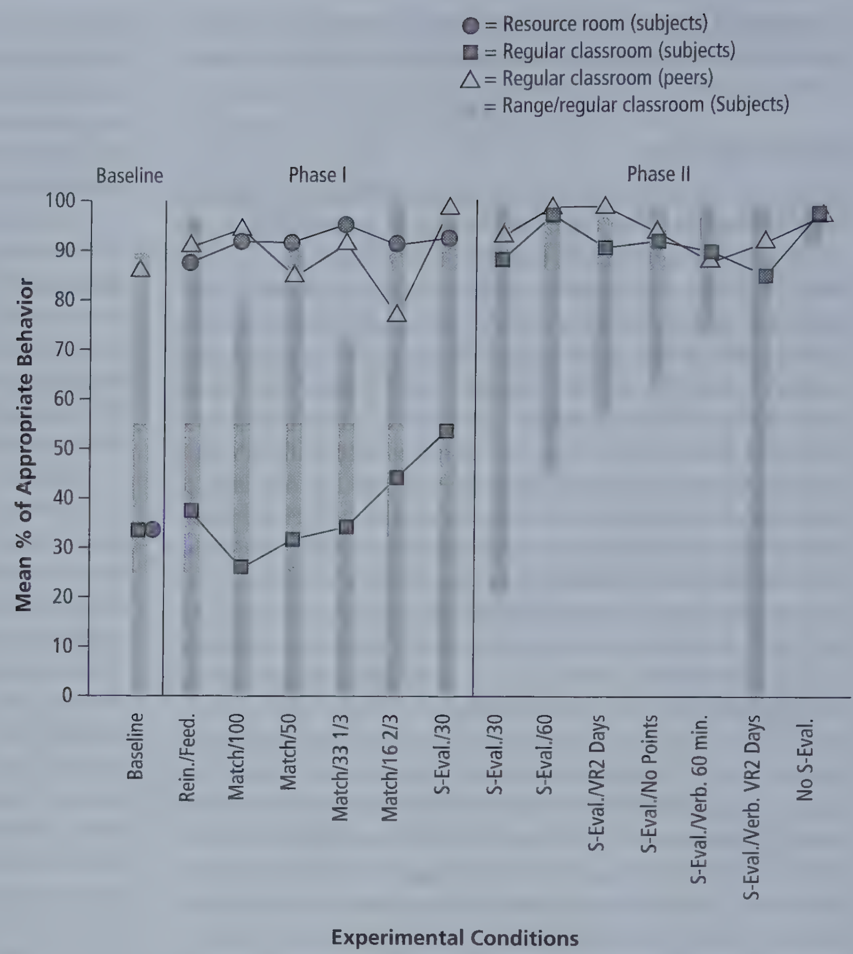


Figure 10.10 Example of using measures of the behavior of a normative sample to assess the social validity of outcomes of a behavior change program.

Based on "Generalization and Maintenance of Treatment Gains of Behaviorally Handicapped Students from Resource Rooms to Regular Classrooms Using Self-Evaluation Procedures" by G. Rhode, D. P. Morgan, and K. R. Young, 1983, *Journal of Applied Behavior Analysis*, *16*, p. 184. Copyright 1984 by the Society for the Experimental Analysis of Behavior, Inc.

Smith, & Lovaas, 1993; Smith, Eikeseth, Klevstrand, & Lovaas, 1997) were key to the recognition and eventual acceptance of early intensive behavior intervention for children with autism as an evidence-based treatment (e.g., Maine Administrators of Services for Children with Disabilities, 2000; Myers, Johnson, & the American Academy of Pediatrics Council on Children with Disabilities, 2007; New York State Department of Health Early Intervention Program, 1999). Subsequent research reporting improvements in scores on standardized tests of adaptive behavior (e.g., *Vineland Adaptive Behavior Scales*, Sparrow, Balla, & Cicchetti, 2005) and child development (e.g., *Bayley Scales of Infant and Toddler Development*, Bayley, 2005) in addition to IQ scores has further strengthened the awareness and widespread acceptance of ABA-informed education and treatment for individuals with autism (e.g., Eikeseth, 2009; Eldevik et al., 2009, 2010; Howard, Sparkman, Cohen, Green, & Stanislaw, 2005).

Scores on tests of academic achievement can also attest to the social validity of behavior change. In their study evaluating the effects of Headsprout Comprehension[8] on the reading comprehension of six elementary students with high-incidence disabilities, Cullen, Alber-Morgan, Schnell, and Wheaton (2014) measured students' performance on comprehension passages based on the Ohio Achievement Assessment, a "high-stakes" test, and on AIMSweb Maze probes. AIMSweb Maze is a norm-referenced curriculum-based measurement (CBM) of reading comprehension (Shinn & Shinn, 2002).

Behavior analysts have created tests that can be used to assess the social validity of some behavior change program outcomes. Iwata, Pace, Kissel, Nau, and Farber (1990) developed the *Self-Injury Trauma Scale (SITS)* to enable researchers and therapists to measure the number, type, severity, and location of injuries produced by self-injurious behavior. The *SITS* yields a Number Index and Severity Index with scores of 0 to 5, and an estimate of current risk. Although the data collected in a treatment program may show significant decreases in the behaviors that produce self-injury (e.g., eye poking, face slapping, head banging), the social significance of the treatment must be validated by evidence of reduced injury. Iwata and colleagues wrote:

> . . . the social relevance of the behavior lies in its traumatic outcome. The measurement of physical injuries prior to treatment can establish the fact that a client or subject actually displays behavior warranting serious attention. . . . Conversely, injury measurement following treatment can corroborate observed changes in behavior because reduction of an injury-producing response below a certain level should be reflected in the eventual disappearance of observable trauma. In both of these instances, data on injuries provide a means of assessing social validity. (pp. 99–100)

Twohig and Woods (2001) used the *SITS* to validate the outcomes of a habit-reversal treatment for the chronic skin

picking of two typically developing adult males. Both men reported that they had engaged in skin picking since childhood, digging the fingernails into the ends of a finger and pulling or scraping the skin, which sometimes caused bleeding, scarring, and infections. Two observers independently rated pretreatment, posttreatment, and follow-up photographs of the two men's hands with the *SITS*. The Number Index (NI) and Severity Index (SI) *SITS* scores on the pretreatment photographs for both men were 1 and 2, respectively, indicating one to four injuries on either hand and distinct but superficial breaks in the skin. NI and SI scores of 0 on the posttreatment photos for both men indicated no apparent injuries. On the follow-up photos taken 4 months after treatment had ended, both men had *SITS* NI and SI scores of 1, indicating red or irritated skin.

Real-World Test

Perhaps the most socially valid way to assess a learner's newly acquired behavior is to put it to an authentic test in the natural environment. For instance, the validity of what three adolescents with learning disabilities had learned about road signs and traffic laws was validated when they passed the Ohio Department of Motor Vehicles test and earned their temporary driver's permits (Test & Heward, 1983). In a similar way, the social validity of the cooking skills being learned by three secondary students with developmental disabilities and visual impairments was tested frequently when their friends arrived at the end of probe sessions to share the food they had just prepared (Trask-Tyler, Grossi, & Heward, 1994). In addition to providing authentic assessments of social validity, real-world tests put the learner's repertoire in contact with naturally occurring contingencies of reinforcement, which may promote maintenance and generalization of the newly acquired behaviors.

Observing desired behavior change in real-world settings may not equate with socially valid improvements in a client's repertoire. This mismatch is most likely to occur when (a) the target behavior differs from the behavior of ultimate value in the natural environment, or (b) the behavior changes are observed during situations in the natural environment that differ from situations when the client's performance of the behavior is most crucial. Stokes, Luiselli, Reed, and Fleming's (2010) study on teaching pass-blocking skills to high school football players provides an example of the first kind of mismatch. The researchers found that participants' performance of a 10-step task analysis for pass-blocking by offensive linemen improved from practice drills to games (see Figure 23.5), but they acknowledged that "the relation between the tackling of a ball carrier in practice and games and performance of the 10 skills remains unknown because we did not measure successful tackles during practice or games" (p. 512).

Directly observing participants' reliable performance of the behavior of ultimate value in real-world settings is no guarantee of socially valid behavior change. If children's ability to use in the natural environment what they had learned in a classroom-based program on safe street crossing behavior was assessed only during days or times when traffic was very light, the results may give a false and overly optimistic impression of the children's ability to safely cross during heavy traffic at rush hour.

A tremendous strength of single-case research designs is transparency: The dependent and independent variables must be described sufficiently to allow implementation by others and to make the data demonstrating behavior changes directly available for anyone to examine. This transparency enables scientific reviewers, practitioners, consumers, and policymakers to judge for themselves the social validity of behavior changes.[9]

EXTERNAL VALIDITY: REPLICATING EXPERIMENTS TO DETERMINE THE GENERALITY OF RESEARCH FINDINGS

External validity refers to the degree to which a functional relation found reliable and socially valid in a given experiment operates under different conditions. An intervention that works only within a circumscribed set of conditions and proves ineffective when any aspect in the original experiment is altered makes a limited contribution to the development of a reliable and useful technology of behavior change. When a carefully controlled experiment has shown that a particular treatment produces consistent and socially significant improvements in the target behavior of a given subject, a series of important questions should then be asked: Will the treatment be as effective if it is applied to other behaviors? Will the procedure continue to work if it is changed in some way (e.g., implemented at a different time of day, by another person, on a different schedule)? Will it work in a setting different from the original experiment? Will it work with participants of different ages, backgrounds, and repertoires? Questions about external validity are neither abstract nor rhetorical; they are empirical questions that if judged important can be addressed by empirical methods.

A functional relation with external validity, or generality, will continue to operate under a variety of conditions. External validity is a matter of degree, not an all-or-nothing property. A functional relation that cannot be reproduced under any conditions other than the exact set of original variables (including the original subject) possesses no external validity. At the other end of the continuum, a procedure that is effective at any time, under any conditions, in any setting, with any behavior, and for any subject has complete generality (a most improbable situation). Most functional relations fall somewhere between the two ends of this continuum, and those found to have higher degrees of generality make the greater contribution to applied behavior analysis. Investigators who use between-groups research methods pursue external validity quite differently than do investigators who use within-subjects research methods.

External Validity and Between-groups Research Design

As stated previously, researchers who use between-groups experimental designs claim two advantages for using large groups of subjects. In addition to the assumption that aggregating the data from individual subjects controls for intersubject variability,[10] group design researchers assume that including many subjects in an experiment increases the external validity of the results. On the surface this assumption is perfectly logical, and it is true when viewed at the proper level of extrapolation.

The more subjects with which a functional relation has been demonstrated, the greater the likelihood it will also be effective with untested subjects who share similar characteristics. And in fact, demonstrating a functional relation with various subjects in different settings is exactly how applied behavior analysts document external validity.

The investigator who claims that the findings of a between-groups study possess generality to other *individuals* in the population from which the experimental subjects were chosen violates a fundamental premise of the between-groups method and ignores a defining characteristic of behavior. The proper inference about the results of a groups-design study is from the *sample* to the *population*, not from the sample to the individual. The random sampling methods in a well-executed between-groups study are followed to ensure that the study's participants exemplify a heterogeneous sample of all relevant characteristics found in the population from which they were selected. The better the sample represents the population from which it is drawn, the less meaningful are the results for any individual subject.

A second problem inherent in extending the results of a between-groups study to other people (and unless care is taken, sometimes even to a subject who participated in the study, as was illustrated in Figure 10.1) is that a between-groups experiment does not demonstrate a functional relation between the behavior of any subject and some aspect of his or her environment. From the perspective of behavior analysis, there is nothing in the results of a between-groups experiment that can have external validity; there is nothing to generalize. Johnston and Pennypacker (2009) made this point repeatedly and well.

> The researcher's primary objective is to obtain data that truly represent the relationship between the independent variable and the dependent variable. If this is not accomplished, nothing else matters. Only when the findings are "true" does the question of the meaningfulness under other circumstances become relevant. (p. 247)
>
> Between groups designs tend to put the cart before the horse. The tactics of exposing different levels of the independent variable to different groups, insuring that these groups contain a large number of participants, and treating their responses collectively provide comparisons that describe no member of any group. By failing to focus on individuals with careful attention to experiment control, these traditional methods greatly decrease the chances of discovering orderly relations in the first place. This makes the question of generality across individuals moot. (p. 346)

Between-groups designs and statistical inference have long dominated research in psychology, education, and other social sciences. Despite its long-standing dominance, the extent to which this research tradition has contributed to an effective technology of behavior change is highly questionable (Baer, 1977; Branch, 2014; Michael, 1974). The field of education is a telling example of the inability of groups-design research to provide data that lead to improved practice (Greer, 1983; Heward & Cooper, 1992). Classroom instruction and curricular decisions are often influenced more by fad, fashion, and ideology than by the cumulative knowledge and understanding provided by rigorous and sustained experimental analysis of the variables of which learning is a function (Heron, Tincani, Peterson, & Miller, 2005; Heward, 2005; Kauffman, 2011; Zane, Weiss, Blanco, Otte, & Southwick, 2015).

Between-groups experimentation is inappropriate for answering the questions of primary interest to applied behavior analysts—empirical questions that can be pursued only by analysis of repeated measures of individual behavior under all relevant conditions. We agree with Johnston and Pennypacker (1993b):

> We find the reasoning underlying all such procedures alien to both the subject matter and the goals of a natural science of behavior and regard the utility of group comparisons as extremely limited, no matter how elegant the mathematical treatment of data they afford. . . . [group comparison experimentation constitutes] a process of scientific inquiry that is almost totally inverted; instead of using questions about natural phenomena to guide decisions about experimental design, models of design are allowed to dictate both the form and content of the questions asked. Not only is this antithetical to the established role of experimentation in science, the types of questions allowed by groups comparison designs are largely inappropriate or irrelevant to gaining an understanding of the determinants of behavior. (pp. 94–95)

Our discussion of the limitations of between-groups designs for behavior analysis should not be confused with a position that group designs and statistical inference have no value as research methods for seeking scientific knowledge about the world. On the contrary, between-groups designs and statistical inference are highly effective tools for seeking answers to the kinds of questions for which they were devised. Properly designed and well-executed between-groups experiments provide answers with a specific degree of confidence (i.e., probability) to questions central to many large-scale evaluations. For example, a government body is less interested in the effects of a new regulation on any individual person (and perhaps even less interested in whether a functional relation exists between the regulation and that person's behavior) than it is in the probability that the behavior of a predictable percentage of the population will be affected by the regulation (see Box 10.1). The former concern is behavioral and best pursued with single-case experimental designs that encompass the methods to address it. The latter concern is actuarial; random sampling and statistical inference of between-groups designs are appropriate tools for addressing it.

External Validity and Applied Behavior Analysis

Behavior analysts assess, establish, and specify the external validity, or *scientific generality* (Branch & Pennypacker, 2013), of single-case research findings by replicating experiments. **Replication** in this context means repeating a previous experiment.[11] Sidman (1960) described two major types of replication—direct and systematic.

BOX 10.1

Between-Groups Research and Applied Behavior Analysis

Between-groups research designs dominate the larger world of behavioral and social sciences. In studies using a between-groups design called the *random control trial* (RCT), each participant is randomly assigned to the treatment group or the no-treatment (control) group. A "double-blind" RCT eliminates potential bias by keeping membership in the two groups hidden from investigators and participants. Various statistical tools are used to determine if pre- and posttreatment measures of the dependent variable differ between the groups. Some policymaking bodies and researchers consider RCTs the gold standard for experimental evidence (e.g., U.S. Department of Education, 2003; Interactive Autism Network, 2019).

Can RCTs and other types of between-groups designs help advance applied behavior analysis, a science whose basic principles and practices for changing behaviors have been discovered, refined, and demonstrated effective by single-case experiments? In our view, between-groups research designs can support applied behavior analysis in two principal ways.

First, results from group studies showing positive effects of behavioral treatments may lead to increased awareness of ABA-informed interventions by policymakers, government officials, education and healthcare administrators who fund treatment initiatives, and consumers who might advocate for evidence-based treatments. "A positive outcome of a group study may not make behavior analysts know more about control of an individual client's behavior, yet such an outcome may nonetheless make more people know about behavior analysis" (Iverson, 2013, p. 26). Although hundreds of single-subject studies had shown the positive effects of early intensive behavioral intervention for children with autism, widespread recognition of the effectiveness of behavioral treatment for autism resulted, in large measure, from group comparison studies (see Eldevik et al., 2009, 2010; Smith, Groen, & Wynn, 2000). In the public health sector, between-groups studies in the United Kingdom demonstrating that the Food Dudes program—a behavioral intervention package that includes peer modeling, video recording, rewards for eating healthy foods, and a parent packet—increased school children's consumption of fruits and vegetables contributed to the Irish Government making the program available to all primary schools in Ireland (Horne et al., 2004, 2009).

Second, disseminating between-groups studies evaluating behavioral interventions in journals and at conferences that do not typically feature single-case design research increases awareness of treatments derived from behavior analysis by scientists in other disciplines, which may lead to collaboration and expansion of research agendas. When behavior analysts publish between-group studies in nonbehavioral journals, as has been the case with transportation safety (Van Houten, Retting, Farmer, Van Houten, & Malenfant, 2000), health care regimens (Raiff, Barry, Ridenour, & Jitnarin, 2016), smoking cessation (Dallery et al., 2017), and drug addiction (Madden, Petry, Badger, & Bickel, 1997), a wider audience learns how the principles and procedures of applied behavior analysis can improve the human condition.

For additional discussion and examples of how group studies can contribute to the science and practice of applied behavior analysis, see Critchfield and Kollins (2001), Critchfield and Reed (2017), and Hanley (2017).

Direct Replication

In a **direct replication**, the researcher makes every effort to duplicate exactly the conditions of an earlier experiment. If the same subject is used in a direct replication, the study is an *intrasubject direct replication*. Intrasubject replication within experiments is a defining characteristic of applied behavior analysis research and the primary tactic for establishing the existence and reliability of a functional relation. An *intersubject direct replication* maintains every aspect of a previous experiment except that different, although similar (e.g., same age, similar repertoires), subjects are involved. Intersubject replication is the primary method for determining the extent to which research findings have generality across subjects.

The many uncontrolled variables in natural settings make the direct replication of experiments outside the laboratory extremely difficult. Nevertheless, intersubject replication is the rule rather than the exception in applied behavior analysis. Although numerous single-case studies involve just one subject (e.g., Ahearn, 2003; Dixon & Falcomata, 2004; Fahmie, Iwata, & Mead, 2016; Tarbox, Williams, & Friman, 2004), the vast majority of published studies in applied behavior analysis include direct intersubject replications. This is because each subject can be considered an intact experiment. For example, a behavior analysis study in which the independent variable is manipulated in exactly the same way for six subjects in the same setting yields five intersubject replications.

Systematic Replication

The direct replication of experiments demonstrates the reliability of a functional relation, but the generality of that finding to other conditions can be established only by a series of experiments in which the conditions of interest are purposefully and systematically varied. In a **systematic replication** the researcher purposefully varies one or more aspects of a previous experiment. When a systematic replication successfully reproduces the results of previous research, it not only demonstrates the reliability of the earlier findings but also adds to the external validity of the earlier findings by showing that the same effect can be obtained under different conditions. In a systematic replication, any aspect of a previous experiment can be altered: subjects, setting, administration of the independent variable, target behaviors.

Because systematic replication can provide new knowledge about the variables under investigation, it offers greater potential rewards than direct replication. Although systematic replication entails some risk, Sidman (1960) described it as a gamble well worth taking.

> If systematic replication fails, the original experiment will still have to be redone, else there is no way of determining whether the failure to replicate stemmed from the introduction of new variables in the second experiment, or whether the control of relevant factors was inadequate in the first experiment.
>
> On the other hand, if systematic replication succeeds, the pay-off is handsome. Not only is the reliability of the original finding increased, but also its generality with respect to other organisms *and* to other experimental procedures is greatly enhanced. Furthermore, additional data are now available which could not have been obtained by a simple repetition of the first experiment. (pp. 111–112)

Sidman went on to explain that economic husbandry of limited resources must also play an important role in the scientist's determination of how a research program should proceed. Direct replication of a long and costly experiment can provide data on only the reliability of a functional relation, whereas systematic replication can provide information on the reliability and generality of the phenomena under investigation.

The external validity of results of between-groups research is viewed as an inherent characteristic of a given experiment, as something that can be directly assessed by examining the methods used to conduct the study (e.g., sampling procedures, statistical tests). But as Birnbrauer (1981) pointed out, external validity is not something a single study *has*, but rather the product of many studies. External validity can be pursued only through the active process of systematic replication.

> Generality is established, or more likely limited, by accumulating studies that are internally valid *and* by placing the results into a systematic context, i.e., seeking out the principles and parameters that particular procedures appear to be enunciating. The most informative studies ask *how* can an earlier positive result be repeated in the present circumstances, with the present problem? (p. 122)

Systematic replications make up much of the applied behavior analysis literature. Indeed, one could argue persuasively that any applied behavior analysis study is a systematic replication of at least some aspect of an earlier experiment. Even when the authors have not pointed it out, virtually every published experiment reveals significant procedural similarity with previous experiments. However, as we are using the term here, systematic replication refers to concerted and directed efforts to establish and specify the generality of a functional relation. For example, Hamlet, Axelrod, and Kuerschner (1984) found a functional relation between demanded eye contact (e.g., "[Name], turn around") and compliance with adult instructions in two 11-year-old school children. Included in the same published report were the results of six replications conducted by the same researchers over a period of 1 year with nine students aged 2 to 21 years. Similar results were reproduced in eight of the nine replication subjects. Although some might consider this an example of direct intersubject replication, Hamlet and colleagues' replications were conducted in various settings (i.e., classrooms, homes, institutions) and therefore should be viewed as a series of systematic replications that demonstrated not only the reliability of the results but also considerable generality across subjects of different ages in different settings.

Systematic replications across subjects sometimes reveal different patterns of effects, which the researcher might then study as a function of specific subject characteristics or contextual variables. For example, Hagopian, Fisher, Sullivan, Acquisto, and LeBlanc (1998) reported the results of a series of systematic replications with 21 inpatient cases of functional communication training with and without extinction and punishment.[12] Lerman, Iwata, Shore, and DeLeon (1997) found that thinning FR 1 schedules of punishment (in which each response is punished) to intermittent punishment produced different effects on the self-injurious behavior of five adults with profound intellectual disabilities.

Some systematic replications are attempts to reproduce the results reported by another researcher in a different situation or context. For example, Saigh and Umar (1983) successfully reproduced in a Sudanese classroom the positive results originally reported with the Good Behavior Game in an American classroom (Barrish, Saunders & Wolf, 1969; see Chapter 28). Saigh and Umar reported that a "considerable degree of support for the cross-cultural utility of the game was established" (p. 343). Subsequent systematic replications of the Good Behavior Game in Northern Ireland (Lynch & Keenan, 2018) and the Netherlands (van Lier, Muthen, van der Sar, & Crijnen, 2004) lend additional support for the cross-cultural generality of the intervention's effectiveness.

Researchers sometimes report multiple experiments, each experiment serving as a systematic replication investigating the variables influencing a given functional relation. For example, Fisher and colleagues (1993) conducted four studies designed to explore the effectiveness of functional communication training (FCT) with and without extinction and punishment.

Systematic replication is evident when a research team pursues a consistent line of related studies over time. Examples of this approach to replication can be found in Neef and colleagues' experiments on the impulsivity of students with attention-deficit hyperactivity disorder (ADHD) (e.g., Bicard & Neef, 2002; Neef, Bicard, & Endo, 2001; Neef, Bicard, Endo, Coury, & Aman, 2005; Neef, Markel, et al., 2005); Miltenberger and colleagues' line of research on teaching safety skills to children (e.g., Himle, Miltenberger, Gatheridge, & Flessner, 2004; Johnson et al., 2006; Knudson et al., 2009; Miltenberger, 2008; Miltenberger et al., 2004; Miltenberger et al., 2005); and Van Houten and colleagues' studies investigating variables affecting motorists' behavior and pedestrian safety spanning four decades (e.g., Huitema, Van Houten, & Manal, 2014; Lebbon, Austin, Van Houten, & Malenfant, 2007; Van Houten, Malenfant, Van Houten, & Retting, 1997; Van Houten, Nau, & Marini, 1980).

The systematic replications necessary to extend a significant line of research often require the independent efforts of investigators at different sites who build upon one another's

work. When independent teams of researchers at different geographical locations report similar findings, the net result is a body of knowledge with significant scientific integrity and technological value. This collective effort speeds and enhances the refinement and rigorous testing of interventions that is necessary to the development and refinement of evidence-based practices (Hitchcock, Kratochwill, & Chezan, 2015; Horner & Kratochwill, 2012).

An example of independent research teams at various sites reporting systematic replications across several decades is a body of studies analyzing the effects of response cards on students' academic engagement, learning, and deportment during group instruction.[13] Investigators have reported a similar pattern of results—increased participation during instruction, improved retention of lesson content, and/or reductions in off-task and disruptive behaviors—with response cards used with a wide range of students (general education students, special education students, English as a second language [ESL] learners), curriculum content (e.g., math, science, social studies, spelling), and instructional settings (e.g., preschool, elementary, middle, secondary, and college classrooms) (e.g., Armendariz & Umbreit, 1999; Bondy & Tincani, 2018; Cakiroglu, 2014; Cavanaugh, Heward, & Donelson, 1996; Davis & O'Neill, 2004; Gardner, Heward, & Grossi, 1994; Hott & Brigham, 2018; Kellum, Carr, & Dozier, 2001; Lambert, Cartledge, Lo, & Heward, 2006; Marmolejo, Wilder, & Bradley, 2004; Wood, Mabry, Kretlow, Lo, & Galloway, 2009).

Recognizing the crucial role replications play in the continued development of applied behavior analysis, the editorial board of the *Journal of Applied Behavior Analysis* in 2017 instituted a new category of articles labeled *Replications*.

> It will be expected that the more direct the replication, the shorter the article. More space, but not much more space, will be needed to describe systematic replications that duplicate procedures and their effects under an ever broadening set of conditions. A lengthy article may indeed be published in *Replications* if, for example, authors duplicate procedures, initially fail to replicate results, but then either demonstrate alternative controlling variables or isolate the conditions under which a functional relation occurs and the conditions under which it does not occur. (Hanley, 2017, p. 4)

Planning, conducting, and presenting the results of an applied behavior analysis study requires a great deal of time, energy, and expertise. Increasingly, organizations representing scientific bodies, practitioner groups, and consumers are seeking systematic presentation of single-case studies. The What Works Clearinghouse, a program within the U.S. Department of Education's Institute of Education Sciences, may have been the first to develop rubrics for evaluating and determining whether single-case studies would be included in research syntheses that they published. Maggin, Briesh, Chafouleas, Ferguson, and Clark (2014) compared five rubrics for identifying empirically supported practices with single-case research. The consensus report from a group of 26 researchers offers guidance to researchers about what to include when submitting a research paper for journal publication using single-case experimental designs (Tate et al., 2016a, 2016b).

EVALUATING APPLIED BEHAVIOR ANALYSIS RESEARCH

Thus far we have identified a considerable number of requirements for good applied behavior analysis. Our purpose now is to summarize those requirements in a sequence of questions one might ask in evaluating the quality of research in applied behavior analysis. Those questions can be organized under four major categories: internal validity, social validity, external validity, and scientific and theoretical significance.

Internal Validity

To determine whether an analysis of behavior has been made, the reader of an applied behavior analysis study must decide whether a functional relation has been demonstrated. This decision requires a close examination of the measurement system, the experimental design, and the degree to which the researcher controlled potential confounds, as well as a careful visual analysis and interpretation of the data.

Definition and Measurement of the Dependent Variable

The initial step in evaluating internal validity is to decide whether to accept the data as valid and accurate measures of the target behavior over the course of the experiment. Some of the important issues to be considered in this decision are captured by the questions shown in Figure 10.11.

Graphic Display

If the data are accepted as a valid and accurate representation of the dependent variable over the course of the experiment, the reader should next assess the extent of stability of the target behavior during each phase of the study. Before evaluating the stability of the data paths, however, the reader should examine the graphic display for any sources of distortion (e.g., scaling of axes, distortions of time on the horizontal axis; see Chapter 6). The researcher or consumer who suspects that any element of the graph may encourage interpretations unwarranted by the data should replot the data using a new set of appropriately scaled axes. In an assessment of the stability of the dependent variable within the different phases of an experiment, the length of the phase or condition must be considered as well as the presence of trends in the data path. The reader should ask whether the conditions in effect during each phase were conducive to practice effects. If so, were these effects allowed to play themselves out before experimental variables were manipulated?

Meaningfulness of Baseline Conditions

The representativeness or fairness of the baseline conditions as the basis for evaluating subsequent performance in the presence of the independent variable should be assessed. In other words, were the baseline conditions meaningful in relation to the target behavior, setting, and research questions addressed by the experiment? For example, consider two experiments by Miller, Hall, and Heward (1995) that evaluated the effects of two procedures for conducting 1-minute time trials during a daily 10-minute practice session on the rate and accuracy with which students

Figure 10.11 Questions that should be asked when evaluating the definition and measurement of the dependent variable in an applied behavior analysis study.

- Was the dependent variable precisely, completely, and unambiguously defined?
- Were examples and nonexamples of the target behavior provided, if doing so would enhance clarity?
- Were the most relevant, measurable dimensions of the target behavior specified (e.g., rate, duration)?
- Were important concomitant behaviors also measured?
- Were the observation and recording procedures appropriate for the target behavior?
- Did the measurement provide valid (i.e., meaningful) data for the problem or research question addressed?
- Was the measurement scale broad and sensitive enough to capture socially significant changes in the target behavior.
- Have the authors provided sufficient information on the training and calibration of observers?
- What procedures were used to assess and ensure the accuracy of measurement?
- Were interobserver agreement (IOA) assessments reported at the levels at which the study's results are presented (e.g., by subject and experimental condition)?
- Were observation sessions scheduled during the times, activities, and places most relevant to the problem or research question?
- Did observations occur often enough and closely enough in time to provide a convincing estimate of behavior change over time?
- Were any contingencies operating in the study that may have influenced the observers' behavior?
- Was there any expectation or indication that the dependent variable may have been reactive to the measurement system? If so, were procedures taken to assess and/or control reactivity?
- Were appropriate accuracy and/or reliability assessments of the data reported?

answered math problems. Throughout all conditions and phases of both experiments, the students were instructed to answer as many problems as they could and they received feedback on their performance. Throughout both experiments, students' worksheets were marked and scored as follows:

> Experimenters marked each student's worksheets by putting an 'X' next to incorrect answers. The number of correct answers over the total number of problems attempted was marked at the top of the first worksheet along with a positive comment to encourage the students to keep trying. If a student's score was lower than his or her highest previous score, comments such as "Keep trying, Sally!" "Work faster!" or "Keep working on it!" were written. Whenever a student achieved his or her highest score to date, comments such as "Great job, Jimmy! This is your best ever!" were written on the packet. In the event a student equaled her highest previous score, "You tied your best score!" was written on the packet.

At the beginning of each session, scored and marked worksheets from the prior day's session were returned to the students. Each session during the 10-minute continuous work period condition that functioned as the baseline condition began with the classroom teachers saying to the students: "I want you to work hard and try to do your best. Answer as many problems as you can. Don't worry if you do not answer all of the problems. There are more problems in the packet than anyone can do. Just try your best" (p. 326).

The initial baseline (A) phase was followed by the two time-trial conditions (B and C) in an A-B-A-B-C-B-C design. Results for students in both classrooms showed a clear functional relation between both of the time-trial conditions and increased correct rate and accuracy over the baseline condition. However, if the classroom teachers had not instructed and reminded the students to do their best and to answer as many problems as they could prior to each baseline session, and if the students had not received feedback on their worksheets, the improved performance during the time-trial conditions would have been suspect. Even if a clear functional relation had been demonstrated against such baseline conditions, applied researchers and consumers could, and should, question the importance of such results. Maybe the children simply did not know they were expected to work fast. Perhaps the students would have solved problems in the baseline condition at the same high rates that they did in the time-trial conditions if they had been told to "go fast" and received feedback on their performance, praise for their improvements, and encouragement to answer more problems. By including the daily instruction to work hard and answer as many problems as they could and by returning the worksheets to the students as components of the baseline condition, Miller and colleagues obtained meaningful data paths during baseline against which to test and compare the effects of the two time-trial conditions.

Experimental Design

The experimental design should be examined to determine the type of experimental reasoning it affords. What elements of the design enable prediction, verification, and replication? Is the design appropriate for the research questions addressed by the study? Does the design effectively control for confounding variables? Does the design provide the basis for component and/or parametric analyses if such questions are warranted?

Brossart, Vannest, Davis, and Patience (2014) introduced the term *design analysis* for assessing single-case experimental designs for threats to internal validity and distinguished design analysis from visual analysis.

> A design analysis examines the structure of the experiment to see if it is valid for asserting a functional relationship, whereas a visual analysis looks at means, trends, immediacy and consistency of change in assessing type and amount of behaviour change. This visual analysis may inform a determination about a functional relationship, but not in isolation from evaluation of the design. (p. 468)

Researcher Concludes Functional Relation Exists	Functional Relation Exists in Nature: Yes	Functional Relation Exists in Nature: No
Yes	Correct Conclusion	Type I Error (false positive)
No	Type II Error (false negative)	Correct Conclusion

Figure 10.12 Ideally, an experimental design and methods of data analysis help a researcher conclude correctly that a functional relation between the independent and dependent variables exists (or does not exist) when in fact such a relation does (or does not) exist in nature. Concluding that the results of an experiment reveal a functional relation when no such relation exists in nature is a Type I error. Conversely, concluding that that independent variable did not have an effect on the dependent variable when such a relation did occur is a Type II error.

Visual Analysis and Interpretation

Although statistical methods for evaluating behavioral data and determining the existence of functional relations in single-subject experimental designs have been proposed for more than 40 years (e.g., Gentile, Roden, & Klein, 1972; Hartmann, 1974; Jones, Vaught, & Weinrott, 1977; Manolov & Solanas, 2013; Pfadt & Wheeler, 1995; Solomon, Howard & Stein, 2015), visual inspection remains the most commonly used, and we believe the most appropriate, method for interpreting data in applied behavior analysis. Chapter 6 describes procedures and guidelines for visual analysis of level, trend, and variability behavioral data within and between experimental conditions. We will briefly present four factors that favor visual analysis in applied behavior analysis.

First, applied behavior analysts have little interest in knowing that a behavior change is a statistically significant outcome of an intervention. Applied behavior analysts are concerned with producing socially significant behavior changes. If a test of statistical inference is required to determine whether an individual client or participant's behavior has changed enough to improve his or her quality of life, the answer is no. Or, as Don Baer noted, "If a problem has been solved, you can *see* that; if you must test for statistical significance, you do not have a solution" (Baer, 1977, p. 171).

Second, visual analysis is well suited for identifying variables that produce strong, large, and reliable effects, which contribute to an effective, robust technology of behavior change. On the other hand, powerful tests of statistical analysis can detect the slightest possible correlation between the independent and dependent variables, which may lead to the inclusion of weak, unreliable variables in the technology.

Two types of errors are possible when determining experimental effect (see Figure 10.12). A **Type I error** (also called a *false positive*) is made when the researcher concludes that the independent variable had an effect on the dependent variable, when in truth no such relation exists in nature. A **Type II error** (also called a *false negative*) is the opposite of a Type I error. In this case, the researcher concludes that an independent variable did not have an effect on the dependent variable, when in truth it did.[14] Ideally, a researcher using well-reasoned experimental tactics coupled with a sound experimental design and buttressed by appropriate methods of data analysis will conclude correctly that a functional relation between the independent and dependent variables exists (or does not exist).

Baer (1977) pointed out that the behavior analyst's reliance on visual inspection to determine experimental effects results in a low incidence of Type I errors but increases the commission of Type II errors. The researcher who relies on tests of statistical significance to determine experimental effects makes many more Type I errors than the behavior analyst, but misses few, if any, variables that might produce some effect.

> Scientists who commit relatively many Type 1 errors are bound to memorize very long lists of variables that are supposed to affect diverse behaviors, some predictable portion of which are not variables at all. By contrast, scientists who commit very few Type 1 errors have relatively short lists of variables to remember. Furthermore, and much more important, it is usually only the very robust, uniformly effective variables that will make their list. Those who will risk Type 1 errors more often will uncover a host of weak variables. Unquestionably, they will know more, although some of that more is wrong, and much of it is tricky. . . . Those who keep their probability of Type 2 errors low do not often reject an actually functional variable, relative to those whose Type 2 error probability is higher. Again, unquestionably, the practitioner with the lower probability of Type 2 errors will know more; but again, the nature of that more is seen often in its weakness, inconsistency of function, or its tight specialization. . . . Individual-subject-design practitioners . . . necessarily fall into very low probabilities of Type 1 errors and very high probabilities of Type 2 errors, relative to their group-paradigm colleagues. As a result, they learn about fewer variables, but these variables are typically more powerful, general, dependable, and—very important—sometimes actionable. These are exactly the variables on which a technology of behavior might be built. (Baer, 1977, pp. 170–171)

A third problem with using statistical methods to determine the existence of functional relations in behavioral data occurs with borderline data sets containing significant amounts of variability. Such data sets should motivate a researcher to engage in additional experimentation in an effort to achieve more consistent experimental control and to discover the factors causing the variability. The researcher who forgoes the additional experimentation in favor of accepting the results of a test of statistical significance as evidence of a functional relation risks leaving important findings in the realm of the unknown. Statistical manipulation of data, no matter how elegant or sophisticated, will never replace manipulation of the variables of which behavior is a function.

> The situation where a significance test might seem helpful is typically one involving sufficient uncontrolled variability in the dependent variable that neither the experimenter nor his readers can be sure that there is an interpretable relationship. This is evidence that the relevant behavior is not under good experimental control, a situation calling for more effective experimentation, not a more complex judgmental aid. (Michael, 1974, p. 650)

Fourth, most statistical tests of significance involve certain assumptions can be applied only to data sets that conform to predetermined criteria (Solomon et al., 2015). If those statistical methods for determining experimental effects were to become highly valued in applied behavior analysis, researchers might begin to design experiments so that such tests could be computed. The resultant loss of flexibility in experimental design would be counterproductive to the continued development of behavioral science.

Visual analysis is not without limitations. Poor interrater agreement on whether various data patterns demonstrate experimental control has long been a concern (e.g., Franklin, Gorman, Beasley, & Allison, 1996; Matyas & Greenwood, 1990; Ottenbacher, 1993). In one of the first studies assessing interrater agreement with visual analysis, DeProspero and Cohen (1979) asked 108 editorial board members and guest reviewers of *JABA* and *Journal of the Experimental Analysis of Behavior (JEAB)* to evaluate hypothetical data displayed in nine ABAB graphs on how convincingly (using a scale of 0 [low] to 100 [high]) each graph demonstrated experimental control. Except for the most obvious graphs, those with "ideal patterns," interrater agreement was relatively low (a 0.61 Pearson correlation coefficient). Diller, Barry, and Gelino (2016) asked 19 *JABA* and *JEAB* editorial board members and 90 Board Certified Behavior Analysts (BCBAs) to evaluate the extent to which pairs of hypothetical data paths displayed in 18 multielement design graphs demonstrated experimental control. These researchers also reported inconsistent agreement among reviewers. In contrast, Kahng and colleagues (2010) found high interrater agreement among 45 editorial board members and associate editors of *JABA* from 2002 to 2004 on the extent to which and whether (yes or no) hypothetical data in 36 ABAB graphs demonstrated experimental control.

A limitation of these studies examining interrater agreement with visual analysis is that participants were provided no context for the data. Without knowledge of crucial contextual factors, an applied behavior analyst cannot interpret graphed data properly. Who was the subject? What behavior was measured? How was it measured? What were the behavior change goals? What was the treatment variable? How, when, where, and by whom was treatment implemented? What scales of behavior change and time do the vertical and horizontal axes represent?

Various structured criteria, formal decision rules, and visual aids for interpretation of single-case data have been proposed (e.g. Fisch, 1998; Fisher, Kelley, & Lomas, 2003; Hagopian et al., 1997; Pfadt, Cohen, Sudhalter, Romanczyk, & Wheeler, 1992). Although multiple raters using the same set of rules or test of statistical aids will—if they apply the rules or compute the math correctly—achieve 100% agreement, that may not necessarily be a desirable outcome. What might be meaningful behavior change in one participant-treatment-setting context may not be in another.

Visual analysis of graphed data has served applied behavior analysis well. Behavior analytic researchers and practitioners will continue to rely on it as the primary means of interpreting behavior change and identifying experimental control. However, statistical analysis of behavioral data can support and promote ABA research in several ways (see Box 10.2), and we encourage behavior analysts to follow the development of statistical tools.

Social Validity

The reader of a published study in applied behavior analysis should judge the social significance of the target behavior, the appropriateness of the procedures, and the social importance of the outcomes (Wolf, 1978).

Chapter 3 detailed many considerations that should guide the applied behavior analyst's selection of target behaviors. The social validity of the dependent variable should be assessed in light of those factors. Ultimately, all of the issues and considerations relative to target behavior selection point to one question: Will an increase (or decrease) in the measured dimension of this behavior improve the person's life directly or indirectly?

The independent variable should be evaluated not only in terms of its effects on the dependent variable, but also in terms of its acceptability, practicality, and cost. Regardless of their effectiveness, treatments perceived by practitioners, parents, and/or clients as unacceptable or undesirable for whatever reason are unlikely to be used. Consequently, such treatments will never have the chance to contribute to a technology of behavior change. The same can be said of independent variables that are extremely complex and thus difficult to learn, teach, and apply. Similarly, treatment procedures that require excessive time and/or money to implement have less social validity than do procedures that can be applied quickly and inexpensively.

Even though behavior change is clearly visible on a graphic display, it may not represent a socially valid improvement for the participant and/or significant others in his environment. In evaluating the results of an applied behavior analysis study, the reader should ask questions such as these: Is the participant (or significant others in the participant's life) better off now that the behavior has changed? Will this new level of performance

BOX 10.2

Supplementing Visual Analysis with Statistical Assessments

When evaluating the effects of behavioral interventions, no amount of statistical computation can replace experimental control of variables and visual analysis of the data. However, visual analysis of data from single-case designs can be supplemented with statistical assessment. Behavioral researchers are using statistical tools with increasing frequency to calculate a standard metric of effect and to calculate the overall effect of multiple studies investigating a common research question.

Effect Size: Quantifying the Magnitude of Treatment Impact

Effect size, a statistical measure of how much a treatment variable influenced a dependent variable, is widely used in group design studies. One type of effect size calculated with *Cohen's d* formula ranges from -1.00 to $+1.00$. Cohen (1988) described effect sizes of $+0.2$ as small, $+0.5$ as medium, and $+0.8$ as large. A negative effect size indicates that a treatment variable has a reductive or harmful effect. Cohen acknowledged the danger of using such terms without noting the study's context and the resources required to obtain the effect (e.g., a small effect from an easy-to-implement treatment accepted by staff and clients may be a more valuable outcome than a large effect obtained with a costly treatment disfavored by staff and clients).

There is no current consensus on how to derive effect sizes for single-case design studies. The U.S. Department of Education's What Works Clearinghouse (2018) recommends calculating the magnitude of treatment effects in single-case studies with a variety of tools, suggesting that stronger conclusions may be drawn if similar results are obtained from multiple approaches. If the magnitude of treatment effects is properly calculated, anyone using the same statistical tools will obtain the same result. The difficulty is "choosing from the bewildering number of statistical methods available" (Brossart, Vannest, Davis, & Patience, 2014, p. 467).

To explore the question of whether different statistical tools can be used for determining the magnitude of effects from single-case studies, Shadish (2014) asked the developers of five different statistical approaches to analyze the nine ABAB data sets from Lambert, Cartledge, Heward, and Lo's (2006) study on the effects of response cards on disruptive behavior during math lessons by nine fourth-grade students (Figure 8.3 shows data for four of the students). The different approaches found effects "not identical but reasonably consistent with each other and with visual inspection" of the data (Shadish, 2014, p. 112). In commenting on these results, Fisher and Lerman (2014) noted that the data sets in Lambert et al.'s study showed "mostly large and unambiguous treatment effects" and that Shadish "acknowledged that more definitive comparison of visual and statistical analyses will require further refinements in statistical methods" (p. 247).

Meta-Analysis: Statistical Synthesis of Related Studies

Meta-analysis entails a variety of statistical techniques for aggregating the effect sizes of a group of studies in a specific area (Borenstein, Hedges, Higgins, & Rothstein, 2009). Two examples are Maggin, O'Keefe, and Johnson's (2011) synthesis of 25 years of special education studies, and Stockard, Wood, Coughin, and Khoury's (2018) quantitative review of 328 studies on Direct Instruction published over a half century.

Meta-analyses that aggregate the individual participant's data across studies are sometimes called *mega-analysis* (H. Cooper & Patall, 2009). An example is Eldevik et al.'s (2009) mega-analysis of 16 group design studies on behavioral intervention for children with autism. The investigators analyzed the results for 453 individual participants (309 children who received behavioral intervention, 39 children who received comparison interventions, and 105 children who were in control groups). Results showed that a far higher percentage of children who received behavioral intervention achieved reliable change in IQ score (at least $+27$ points) and adaptive behavior (at least 21 points) than did children in the comparison and control groups.

Readers interested in learning more about statistical tools for single-case data can examine the What Works Clearinghouse (2018) *Standards Handbook*; explore the free web-based calculators and manuscripts archived by Vannest, Parker, Gonen, and Adiguzel (2016); and study articles by single-case researchers with expertise in statistical methods (e.g., Fisher & Lerman, 2014; Kratochwill & Levin, 2014; Solomon, Howard, & Stein, 2015).

result in increased reinforcement (or decreased punishment) for the subject now or in the future? In some instances it is relevant to ask whether the subject (or significant others) believes that her behavior has improved (Wolf, 1978).

Maintenance and Generalization of Behavior Change

Improvements in behavior are most beneficial when they are long-lasting, appear in other appropriate environments, and spill over to other related behaviors. Producing these kinds of effects is a major goal of applied behavior analysis. (Chapter 30 examines strategies and tactics that promote generalized behavior change.) When evaluating applied behavior analysis research, consumers should consider the maintenance and generalization of behavior change in their evaluation of a study. An impressive behavior change that does not last or is restricted to a specialized training setting has limited value.

Did the researchers report the results of assessing maintenance and generalization through follow-up observations and

measurement in nontraining environments? If maintenance and/or generalization were not evident in follow-up observations, did the experimenters modify their design and implement procedures in an attempt to produce and analyze the occurrence of maintenance and/or generalization? Additionally, the reader should ask whether response generalization—changes in functionally similar but untreated behaviors concomitant with changes in the target behavior(s)—is an appropriate concern in a given study. If so, did the experimenters attempt to assess, analyze, or discuss this phenomenon?

External Validity

As discussed earlier, the generality of an experiment's findings to other subjects, settings, and behaviors cannot be assessed by inherent aspects of the study itself. The generality of a behavior–environment relation can be established only through the active process of systematic replication. Therefore, the reader of an applied behavior analysis study should compare the study's results with those of other published studies sharing relevant features. The authors of a published report identify in the paper's introduction the experiments that they believe are most relevant. To judge whether and the extent to which a given study contributes to the generality of previous findings, the reader must examine previous studies in the literature and compare the results of these studies with those of the current experiment.

Although external validity should not be considered a characteristic of a study per se, various features of a single-case experiment suggest an expected, or likely, level of generality for the results. For example, an experiment that demonstrated a functional relation of similar form and degree in six subjects of different ages, cultural and linguistic backgrounds, and skills levels suggests a higher probability of generality to other subjects than would an identical study demonstrating the same results across six subjects of the same age, background, and skill level.[15] Additional confidence in the generality of the results may be warranted if the experiment was conducted in various settings and different people administered the independent variable.

A research finding without extensive external validity can make a meaningful contribution to scientific knowledge. Johnston and Pennypacker (2009) point out, "A procedure can be quite valuable even though it is effective in a narrow range of conditions, as long as we know what those conditions are" (pp. 343–344).

Theoretical Significance and Conceptual Sense

A published experiment should also be evaluated in terms of its scientific merit. It is possible for a study to demonstrate a functional relation between the independent variable and a socially important target behavior—and thus be judged significant from an applied perspective—yet contribute little to the advancement of the field.[16] It is possible to reliably reproduce an important behavior change while at the same time not fully understand which variables are responsible for the observed functional relation. Sidman (1960) differentiated this kind of simple reliability from "knowledgeable reproducibility," a more complete level of analysis in which all of the important factors have been identified and controlled.

After examining the majority of experimental articles published in the first 10 volumes of the *Journal of Applied Behavior Analysis* (1968 to 1977), Hayes, Rincover, and Solnick (1980) concluded that a technical drift had occurred in the field away from conceptual analyses and toward an emphasis on client cure. They warned of a likely loss of scientific understanding as a result of focusing purely on the technical aspects of improving behavior in applied settings, and they recommended an increased effort to perform more thorough analyses of behavior.

> [T]hese more complicated analyses may increase our knowledge of the actual functional variables and subsequently increase our ability to generate more efficient and general behavioral programs. Perhaps, we have gone too far in our attempt to be *immediately* applied at the expense of being *ultimately* more effective, in failing to encourage more analogue and analytical studies that have treatment implications. (Hayes et al., 1980, p. 283)

Baer, Wolf, and Risley (1987), writing in the 20th anniversary issue of the *Journal of Applied Behavior Analysis*, emphasized the need to shift from demonstrations of behavior changes—as convincing as they might be—to a more complete analysis and conceptual understanding of the principles that underlie the successful demonstrations.

> Twenty years ago, *analytic* meant a convincing experimental design, and *conceptual* meant relevance to a comprehensive theory about behavior. . . . Now, applied behavior analysis is more often considered an analytic discipline only when it demonstrates convincingly how to make specified behavior changes *and* when its behavior-change methods make systematic, conceptual sense. In the past 20 years, we have sometimes demonstrated convincingly that we had changed behavior as specified, but by methods that did not make systematic, conceptual sense—it was not clear *why* those methods had worked. Such cases let us see that we were sometimes convincingly applied and behavioral, yet even so, not sufficiently analytic. (p. 318)

Morris (1991) expressed concern over the rise of research that "demonstrates the effects of behavioral interventions at the expense of discovering . . . actual controlling relationships" (p. 413). Numerous authors have discussed the importance of focusing on the analysis side of applied behavior analysis as much as the applied side (e.g., Baer, 1991; Birnbrauer, 1979, 1981; Deitz, 1982; Iwata, 1991; Johnston, 1991; see Schlinger, 2017 for a review).

We agree with the need for more sophisticated, thorough analyses of the variables controlling socially important behavior. Fortunately, examination of the recent literature reveals numerous examples of studies featuring component analyses and parametric analyses that are necessary steps to a more complete understanding of behavior—an understanding that is prerequisite to the development of a thoroughly effective technology of behavior change. Several of the studies cited earlier in this chapter as examples of systematic replication incorporated component and parametric analyses.

The extent of a phenomenon's generality is known only when all of the necessary and sufficient conditions for its

reproducibility have been specified. Only when all of the variables influencing a functional relation have been identified and accounted for can an analysis be considered complete. Even then, the notion of a complete analysis is misleading: "Further dissection or elaboration of either variable in a functional relation inevitably reveals fresh variability, and analysis proceeds anew. . . . the analysis of behavior can never be complete" (Pennypacker, 1981, p. 159).

Evaluation of scientific significance takes into consideration such things as the authors' technological description of the experiment as well as their interpretation and discussion of the results. Are the procedures described in sufficient detail so that at least the unique aspects of the study can be replicated?[17]

Readers should consider the level of conceptual integrity displayed in an experimental report. Does the literature review reveal a careful integration of the study with previous research? Does the literature review provide sufficient justification for the study's research questions? Are the authors' conclusions based on the data obtained in the study? Have the authors respected the difference between basic principles of behavior and behavior change tactics? Do the authors speculate beyond the data without making it clear that they are doing so? Do the authors suggest directions for additional research to further analyze the problem studied? Is the study important for reasons other than the results actually obtained? For example, an experiment that demonstrates a new measurement technique, investigates a new dependent or independent variable, or incorporates a novel tactic for controlling a confounding variable can contribute to the scientific advancement of behavior analysis, even though the study failed to achieve experimental control or produce socially significant behavior change.

Numerous criteria and considerations are involved in evaluating the "goodness" of an applied behavior analysis study. Although each criterion is important on one level or another, it is unlikely that any experiment will meet all of the criteria. And, in fact, it is unnecessary for an experiment to do so to be considered worthwhile. Nevertheless, incorporating these considerations into a study enhances its social significance and scientific value.

SUMMARY

Importance of the Individual Subject in Behavior Analysis Research

1. Behavior analysis research methods feature repeated measures of the behavior of individual organisms within and across experimental conditions.
2. By contrast, a between-groups experiment follows a sequence like this:
   - A pool of subjects is selected randomly from the population of interest.
   - The subjects are divided randomly into an experimental group and a control group.
   - A pretest measure of the dependent variable is obtained for all subjects in each group, individual pretest scores for the subjects in each group are combined, and the means and standard deviations are calculated.
   - Subjects in the experimental group are exposed to the independent variable.
   - A posttest measure of the dependent variable is obtained for all subjects, and means and standard deviations are computed for each group.
   - Various statistical tests compare any changes in each group's scores from pretest to posttest, applying to the data that enable inferences regarding the likelihood that any differences between the two groups' performances can be attributed to the independent variable.
3. Knowing that the average performance of a group of subjects changed may not reveal anything about the performance of individual subjects.
4. To be most useful, a treatment must be understood at the level at which people come into contact with it and are affected by it: the individual level.
5. Mean performance of a group of subjects hides intersubject variability in the data. Attempting to cancel out variability through statistical manipulation neither eliminates its presence in the data nor controls the variables responsible for it.
6. When repeated measurement reveals significant variability, the researcher should seek to identify and control the factors responsible for it.
7. The researcher who attributes the effects of unknown or uncontrolled variables to chance is unlikely to identify and analyze important variables.
8. To control the effects of any variable, a researcher must either hold it constant throughout the experiment or manipulate it as an independent variable.
9. Group data may not represent actual behavioral processes.
10. A great strength of within-subject experimental designs is the convincing demonstration of a functional relation made possible by replication within the design itself.
11. The overall performance of a group is socially significant in many situations.
12. When group results do not represent individual performances, researchers should supplement group data with individual results.
13. When the behavior analyst cannot control access to the experimental setting or identify individual subjects, the dependent variable must consist of the responses made by individuals who enter the experimental setting.

Importance of Flexibility in Experimental Design

14. An effective experimental design is any sequence of independent variable manipulations that produces data that effectively and convincingly address the research question(s).

15. No ready-made experimental designs await selection, nor is there a set of rules that must be followed.

16. To investigate the research question(s) of interest, an experimenter must often build an experimental design that employs a combination of analytic tactics.

17. A component analysis is any experiment designed to identify the active elements of a treatment condition, the relative contributions of different components in a treatment package, and/or the necessity and sufficiency of treatment components. Two basic methods are used for conducting component analyses:
    - drop-out component analysis: the investigator presents and assesses the treatment package as a whole, then systematically removes components.
    - add-in component analysis: the investigator presents and assesses components individually or in combination, then presents the complete treatment package.

18. The most effective experimental designs use ongoing evaluation of data from individual subjects as the basis for employing the three elements of baseline logic—prediction, verification, and replication.

Internal Validity: Controlling Potential Sources of Confounding in Experimental Design

19. Experiments that demonstrate a clear functional relation between the independent variable and the target behavior have a high degree of internal validity.

20. An experimental design's strength is determined by the extent to which it (a) demonstrates a reliable effect and (b) eliminates or reduces the likelihood that factors other than the independent variable produced the behavior change.

21. Although the term *experimental control* implies that the researcher controls the subject's behavior, the researcher can control only some aspect of the subject's environment.

22. A confounding variable is an uncontrolled factor known or suspected to have exerted influence on the dependent variable.

23. Confounding variables can be viewed as related primarily to one of four elements of an experiment: subject, setting, measurement of the dependent variable, and independent variable.

24. A placebo control separates effects that may be produced by a subject's expectations of improvement as a result of receiving treatment from the effects actually produced by the treatment.

25. With a double-blind control procedure neither the subject(s) nor the observers know when the independent variable is present or absent.

26. Treatment integrity refers to the extent to which the independent variable is implemented or carried out as planned.

27. Procedural fidelity refers to the extent to which procedures in all conditions of an experiment, including baseline, are correctly implemented.

28. Low treatment integrity invites a major source of confounding into an experiment, making it difficult, if not impossible, to interpret the results with confidence.

29. One threat to treatment integrity, treatment drift, occurs when application of the independent variable during later phases of an experiment differs from the way the treatment was applied at the study's outset.

30. Achieving a high level of treatment integrity begins with a precise operational definition of treatment procedures.

31. Simplifying and standardizing the independent variable and providing criterion-based training and practice for the people responsible for implementing it enhance treatment integrity.

32. Researchers should not assume that a person's general competence or experience in the experimental setting, or that providing the intervention agent with detailed written instructions or a script, will ensure a high degree of treatment integrity.

33. Treatment integrity/procedural fidelity data describe the extent to which the actual implementation of experimental procedures matches their descriptions in the method section of a research report.

34. Treatment integrity/procedural fidelity data are required for a study's inclusion in systematic reviews of literature conducted by professional and government organizations.

Social Validity: Assessing the Applied Value of Behavior Changes and the Treatments That Accomplish Them

35. The social validity of an applied behavior analysis study should be assessed in three ways: the social significance of the behavior change goals, the appropriateness of the intervention, and the social importance of the results.

36. Experts and persons who use the prospective skill in natural environments can help researchers identify socially valid target behaviors.

37. Socially valid goals can be determined empirically by assessing the performance of individuals judged to be highly competent, and experimentally manipulating different levels of performance to determine socially valid outcomes.

38. Several scales and questionnaires for obtaining consumers' opinions of the acceptability of behavioral interventions have been developed.

39. Preference for interventions can be assessed directly by arranging experimental conditions in which practitioners or clients can choose which of multiple treatments to implement or receive.

40. Methods for assessing the social validity of outcomes include: (a) having consumers rate the social validity of participants' performance, (b) asking experts to evaluate participants' performance, (c) comparing participants' performance to that of a normative sample, (d) standardized assessment instruments, and (e) real-world tests of participants' performance.

External Validity: Replicating Experiments to Determine the Generality of Research Findings

41. External validity refers to the degree to which a functional relation found reliable and socially valid in a given experiment operates under different conditions.

42. Group design researchers assume that including many subjects increases the external validity, or generality, of a study's findings to other individuals.

43. The assumption of increased generality violates a fundamental premise of the between-groups method: The proper inferences about the results of a groups design study are from the sample to the population, not from the sample to the individual.

44. A between-groups experiment does not demonstrate a functional relation between the behavior of any subject and some aspect of his or her environment; the external validity of its findings is moot.

45. Although between-groups designs and tests of statistical significance are necessary and effective tools for answering certain types of research questions, they have contributed little to an effective technology of behavior change.

46. The generality of research findings in applied behavior analysis is assessed, established, and specified by replicating experiments.

47. In a direct replication, the researcher makes every effort to duplicate exactly the conditions of an earlier experiment.

48. In a systematic replication, the researcher purposefully varies one or more aspects of an earlier experiment.

49. When a systematic replication successfully reproduces the results of previous research, it not only demonstrates the reliability of the earlier findings but also adds to the generality of the earlier findings by showing that the same effect can be obtained under different conditions.

50. Systematic replications occur in both planned and unplanned ways through the work of many experimenters in a given area, and they result in a body of knowledge with significant scientific integrity and technological value.

Evaluating Applied Behavior Analysis Research

51. The quality and value of an applied behavior analysis study may be evaluated by seeking answers to a sequence of questions related to internal validity, social validity, maintenance and generalization of behavior change, external validity, and scientific and theoretical significance.

52. A Type I error occurs when a researcher concludes that the independent variable had an effect on the dependent variable when it did not. A Type II error occurs when a researcher concludes that the independent variable did not have an effect on the dependent variable when it did.

53. Visual analysis effectively identifies variables that produce strong, large, and reliable effects, which contribute to an effective and robust technology of behavior change. Statistical analysis detects the slightest possible correlations between the independent and dependent variables, which may lead to the identification and inclusion of weak and unreliable variables in the technology.

54. A study can yield significant results from an applied perspective—demonstrate a functional relation between the independent variable and a socially important target behavior—yet contribute little to the advancement of the field.

55. Only when all of the variables influencing a functional relation have been identified and accounted for can an analysis be considered complete.

56. When evaluating the scientific significance of a research report, readers should consider the technological description of the experiment, the interpretation and discussion of the results, and the level of conceptual sense and integrity.

KEY TERMS

add-in component analysis
component analysis
direct replication
double-blind control
drop-out component analysis
placebo control
procedural fidelity
replication
social validity
systematic replication
treatment drift
treatment integrity
treatment package
Type I error
Type II error

MULTIPLE-CHOICE QUESTIONS

1. The most useful level to understand the effect of a treatment is the:
   a. Group level
   b. Individual level
   c. Procedural level
   d. Implementation level
   Hint: (See "Importance of the Individual Subject in Behavior Analysis Research")

2. Applied behavior analysis research has identified effective interventions for socially significant behavior by focusing on the
   a. Behavior of individuals with cognitive delays
   b. Common behavior of groups of individuals
   c. Behavior of individual subjects
   d. Behavior of professionals
   Hint: (See "Importance of the Individual Subject in Behavior Analysis Research")

3. Combining analytic tactics in an experimental design is a good idea when
   a. It allows for a more convincing demonstration of experimental effect
   b. The research question is too broad for a single design
   c. The procedures are quite complex
   d. None of these is a justification for combining tactics
   Hint: (See "Experimental Designs Combining Analytic Tactics")

4. The research tactic for examining two or more elements of a treatment package is called a(n)
   a. Alternating treatment design
   b. Multiple baseline across interventions
   c. Multiple treatments
   d. Component analysis
   Hint: (See "Experimental Designs Combining Analytic Tactics")

5. Experiments that demonstrate a clear functional relationship between the independent variable and the observed behavior are said to have a high degree of
   a. Internal validity
   b. External validity
   c. Social validity
   d. Reliability
   Hint: (See "Internal Validity: Controlling Potential Sources of Confounding in Experimental Design")

6. An uncontrolled factor that is known or suspected to have exerted influence on the dependent variable is called a(n)
   a. Extraneous factor
   b. Confounding variable
   c. Behavioral control
   d. Maintaining event
   Hint: (See "Internal Validity: Controlling Potential Sources of Confounding in Experimental Design")

7. All of the following are methods for assessing the social validity of outcomes except:
   a. Comparing participants' performance to the performance of a normative sample
   b. Using standardized assessment instruments
   c. Asking consumers to rate the social validity of participant outcomes
   d. Asking the researchers to rate the social validity of participant outcomes
   Hint: (See "Social Validity: Assessing the Applied Value of Behavior Changes and the Treatments That Accomplish Them")

8. The ultimate purpose of social validity assessment is to
   a. Guide behavior change program development and application
   b. Improve researchers' chances of publication
   c. Improve the relationship between researchers and clients
   d. Increase the magnitude of behavior change
   Hint: (See "Social Validity: Assessing the Applied Value of Behavior Changes and the Treatments That Accomplish Them")

9. The degree to which a functional relationship can be demonstrated in circumstances that are different from the experimental conditions is called:
   a. Internal validity
   b. Treatment acceptability
   c. Replication
   d. External validity
   Hint: (See "External Validity: Replicating Experiments to Determine the Generality of Research Findings")

10. Which of the following research designs have contributed the least in terms of effective technology for behavior change?
    a. Group comparison designs
    b. Multiple baseline across settings
    c. Reversal designs
    d. Component analyses
    Hint: (See "External Validity: Replicating Experiments to Determine the Generality of Research Findings")

11. Which of the following is a factor not to consider when determining whether or not a functional relation has been demonstrated in an experiment?
    a. The degree to which the researcher controlled potential confounds
    b. An examination of the measurement system
    c. Visual analysis and interpretation of the data
    d. Systematic replication of effects

    Hint: (See "Evaluating Applied Behavior Analysis Research")

12. Which of the following is not a reason that applied behavior analysts favor visual analysis of data over statistical methods?
    a. Statistical significance is not the target outcome for behavior analysts.
    b. Visual analysis is well suited for demonstrating strong effects.
    c. Highly variable data may not be statistically significant.
    d. Predetermined criteria of statistical tests offer less flexibility in experimental design.

    Hint: (See "Evaluating Applied Behavior Analysis Research")

ESSAY-TYPE QUESTIONS

1. Give at least two reasons why knowledge of group performance means is inadequate information for the behavior analyst to evaluate the effectiveness of a treatment.

   Hint: (See "Group Data May Not Represent the Performance of Individual Subjects")

2. Explain two benefits of building an experimental design using a combination of analytic tactics.

   Hint: (See "Experimental Designs Combining Analytic Tactics")

3. Define treatment integrity and explain its importance in applied behavior analysis research.

   Hint: (See "Independent Variable Confounds")

4. Describe the ways that Wolf (1978) suggested that social validity should be assessed in applied behavior analysis.

   Hint: (See "Social Validity: Assessing the Value of Behavior Changes and the Treatments That Accomplish Them")

5. Compare and contrast direct replication with systematic replication. What does each type of replication demonstrate?

   Hint: (See "External Validity and Applied Behavior Analysis")

6. Explain the difference between a Type I and a Type II error in experimental research.

   Hint: (See "Evaluating Applied Behavior Analysis Research")

NOTES

1. Between-groups experimental designs are characterized by few measures of the dependent variable (often just two: pretest and posttest) across many subjects. In contrast, single-case studies require many measures of the dependent variable across a few participants (sometimes just one). Behavioral researchers occasionally employ pretest-treatment-posttest designs, typically reporting the results for each subject as well as group means (e.g., Keintz, Miguel, Kao, & Finn, 2011; Sawyer et al., 2017).
2. This brief sketch of the simplest form of a between-groups design study omits many important details and controls. Readers interested in a thorough explication of group research methods should consult an authoritative text (e.g., Adams & Lawrence, 2019; Campbell and Stanley, 1963; Shaddish Cook, & Campbell, 2001). Blampied (2013), Branch (2014), and Colquhoun (2014, 2016) describe how the decades-long reliance on statistical significance testing, first championed by the agricultural scientist and geneticist R. A. Fisher (1925, 1935), has impaired robust and reliable scientific progress.
3. A complete discussion of the problems posed by commingling data from multiple subjects is beyond the scope of this text. We encourage readers seeking a more complete understanding of these important issues to read Branch and Pennypacker (2013), Johnston and Pennypacker (1980, 1993b, 2009), and Sidman (1960).
4. The posttest data point in the left-hand graph of Figure 10.1 is reminiscent of the man whose bare feet were in a bucket of ice while his head was on fire. Asked how he was feeling, the man replied, "On the average, I feel fine."
5. The analytic tactics presented in Chapters 8 and 9 should not be considered experimental designs in and of themselves for another reason: All experiments incorporate design elements in addition to the type and sequence of independent variable manipulations (e.g., subjects, setting, dependent variable, measurement system).
6. Time-out and differential reinforcement of other behavior (DRO) are described in Chapters 15 and 25, respectively.
7. Treatment integrity can be manipulated as an independent variable to analyze the effects of full versus various levels of fidelity, or types of treatment "mistakes," in implementing an intervention on participants' behavior (e.g., Carroll, Kodak, & Fisher, 2013; DiGennaro Reed, Reed, Baez, & Maguire, 2011; Groskreutz, Groskreutz, & Higbee, 2011; Pence & St. Peter, 2015; St. Peter, Byrd, Pence, & Foreman, 2016 [see discussion of omission errors and commission errors in Chapter 15]; Wilder, Atwell, & Wine, 2006 [see Figure 8.13]).
8. Headsprout Comprehension and Headsprout Early Reading are computer-based reading programs for beginning readers and were developed by behavior analysts (Layng, Twyman, & Stikeleather, 2004; Twyman, Layng, Stikeleather, & Hobbins, 2005).
9. Additional discussions of social validity and procedures for assessing it can be found in Carroll and St. Peter (2014); Common and Lane (2017); Snodgrass, Chung, Meadan, and Halle (2018); and the special section on social validity in the Summer 1991 issue of the *Journal of Applied Behavior Analysis*.

10. Assuming group data describe the behavior of individual participants poses little problem when intersubject differences are negligible or of no particular interest (Blampied, 2013). One kernel of corn is genetically and functionally identical to the millions of other kernels harvested from the same acre of farmland. People are quite another matter.
11. Johnston and Pennypacker (1980) pointed out a distinction between replicating an experiment and reproducing its results. They stated that the quality of the replication should be judged only by "the extent to which equivalent environmental manipulations associated with [the original experiment] are duplicated. . . . Thus, one replicates procedures in an effort to reproduce effects" (pp. 303–304). However, when most researchers report a "failure to replicate," they mean that the results of the replication did not match those obtained in the earlier research (e.g., Ecott, Foate, Taylor, & Critchfield, 1999; Friedling & O'Leary, 1979).
12. Functional communication training (FCT) is described in Chapter 26.
13. Response cards are small cards, signs, or items that students hold up to display their answers to teacher-posed questions or problems.
14. The boy who cried wolf caused both Type I and Type II errors. First everyone believed there was a wolf, when there wasn't. Next everyone believed there was no wolf, when there was. Substitute "effect" for "wolf" and you'll never confuse Type 1 and II errors again (Stafford, 2015).
15. The American Psychological Association (2010) recommends that authors indicate the sex, disability status, socio-economic status, and ethnicity of research participants. Li, Wallace, Ehrhardt, and Poling (2017) found that only 4 of 288 intervention articles published in behavior analysis journals for 2013 through 2015 included information on all four descriptors.
16. It is important to remember that although some research in applied behavior analysis can rightfully be criticized as superficial because it adds little to our conceptual understanding of behavior, studies in which meaningful target behaviors are improved to a socially valid level by the application of a socially valid treatment variable (whether a package or not) are never superficial to the participants and the significant others who share their environment.
17. Ideally, published procedural descriptions should include sufficient detail to allow an experienced investigator to replicate the experiment. However, space limitations of most journals often prohibit such detail. The common and recommended practice in replicating published studies is to request complete experimental protocols from the original investigator(s).

PART 4

Reinforcement

The three chapters in Part 4 are devoted to reinforcement, the most important and widely applied principle of behavior analysis. Reinforcement, a deceptively simple behavior–consequence relation, is the fundamental building block of operant behavior. Chapter 11 presents the operation and defining effect of reinforcement, explains how reinforcers are classified and identified, describes briefly how antecedent stimulus conditions modulate the effects of reinforcement, discusses factors that influence the effectiveness of reinforcement, outlines experimental control techniques for verifying whether a positive reinforcement contingency is responsible for increased responding, and offers guidelines for using reinforcement effectively.

In Chapter 12, Richard Smith and Brian Iwata describe one of the most consistently misunderstood principles of behavior. Smith and Iwata define negative reinforcement, an operant contingency in which responding increases as a result of the termination, reduction, or postponement of a stimulus as a behavioral consequence; compare and contrast it with positive reinforcement and punishment; distinguish between escape and avoidance contingencies; describe events that may serve as negative reinforcers; illustrate how the application of negative reinforcement can be used to strengthen desired behavior; and discuss ethical issues that arise when using negative reinforcement.

One of Skinner's most important discoveries was that reinforcement does not have to follow every response. In fact, under many intermittent schedules of reinforcement—in which reinforcement follows some, but not all, occurrences of the target behavior—responding occurs at higher, more consistent, rates than under a continuous schedule of reinforcement, in which each response is reinforced. Chapter 13 describes how intermittent reinforcement can be scheduled based on various combinations of response and/or temporal requirements and identifies characteristic patterns of responding associated with each schedule. Practitioners who understand how different schedules of reinforcement influence behavior can program reinforcement for more effective and efficient acquisition of new skills and for improved performance and endurance of established skills.

CHAPTER 11

Positive Reinforcement

LEARNING OBJECTIVES

- Define and provide examples of positive reinforcement.
- Define and provide examples of conditioned and unconditioned reinforcement.
- Describe and provide examples of the operant conditioning paradigm (i.e., the three-term and four-term contingencies).
- Identify potential reinforcers.
- Use appropriate parameters and schedules of reinforcement to identify reinforcers.
- Use response-deprivation procedures (e.g., the Premack principle).
- Identify control procedures for positive reinforcement.
- Use positive reinforcement effectively.

In looking back, it seems to me that the most important thing I learned in graduate school was from another student, Burrhus Frederic Skinner (I called him Burrhus, others called him Fred). This man had a box, within which was a smaller box, within which he would place a hungry laboratory rat. When the animal, in its explorations, would depress a lever that projected from one wall, a pellet of food would be discharged into a tray beneath the lever. Under such conditions, the rat would learn, in a matter of minutes, sometimes seconds, how to get its meal by depression of the lever. It would even keep on pressing, sometimes at a rapid rate, when pellets were delivered only now and then; and if the food supply was cut off entirely, the animal would still keep working for awhile.

—Fred Keller (1982, p. 7)

Positive reinforcement is the most important and most widely applied principle of behavior. Although some people continue to believe that the results of laboratory research with animals is not applicable to human learning, by the mid-1960s researchers established the significant role of positive reinforcement in education and treatment. "It is safe to say that without Skinner's detailed laboratory analyses of reinforcement (Skinner, 1938), there would be no field of 'applied behavior analysis' today, least not as we know it" (Vollmer & Hackenberg, 2001, p. 241).

Fittingly, the lead article in the first issue of the *Journal of Applied Behavior Analysis* reported an experiment showing the effects of positive reinforcement on student behavior (Hall, Lund, & Jackson, 1968). Six elementary students who were disruptive or dawdled frequently participated in this classic study. The dependent variable, study behavior, was defined individually for each student, depending on the subject matter being taught, but it generally consisted of the student being seated and oriented toward the appropriate object or person (e.g., looking at course materials or the lecturing teacher) and class participation (e.g., writing the assignment, answering the teacher's question). The independent variable was teacher attention, cued by an observer who held up a small square of colored paper not likely to be noticed by the target student. On this signal, the teacher attended to the child by moving to his desk and making a verbal comment or giving him a pat on the shoulder.

The effects of contingent teacher attention on the behavior of all six students were striking. Figure 11.1 shows the results for Robbie, a third-grader who was "a particularly disruptive student who studied very little" (p. 3). During baseline, Robbie engaged in study behavior for an average of 25% of the observed intervals. The remainder of the time he snapped rubber bands, played with objects in his pocket, talked and laughed with classmates, and played with an empty milk carton from his earlier-served drink. The majority of attention Robbie received followed nonstudy, disruptive behaviors.

Following baseline, the experimenters showed the teacher a graph of Robbie's study behavior, presented the results of previous studies in which contingent adult attention had improved child behavior, and discussed the fundamentals of providing contingent social reinforcement.

The results of contingent reinforcement were as follows: During Reinforcement 1, Robbie's study behavior increased to a mean of 71%. When a reversal to baseline conditions was introduced, his study behavior decreased to a mean of 50%; but when Robbie's teacher again provided contingent attention for study behavior (Reinforcement 2), his study behavior recovered and stabilized at a level ranging between 70% and 80% of the observed intervals. Followup observations over a 14-week period showed that Robbie's study behavior had maintained at 79%. The teacher reported positive behavior changes associated with Robbie's increased study behavior. By the final week of

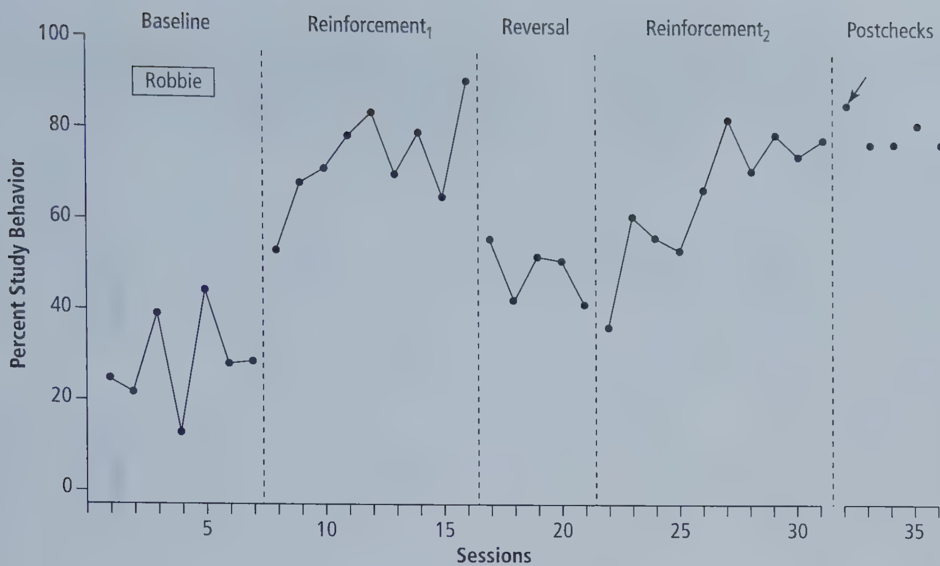


Figure 11.1 Percentage of intervals of study behavior by a third-grade student during baseline and reinforcement conditions. The arrow to the first postcheck data point shows when cueing the teacher to provide attention was discontinued.

Based on "Effects of Teacher Attention on Study Behavior" by R. V. Hall, D. Lund, and D. Jackson, 1968, *Journal of Applied Behavior Analysis, 1*, p. 3. Copyright 1968 by the Society for the Experimental Analysis of Behavior, Inc.

Reinforcement 2, Robbie completed his spelling assignments more consistently, his disruptive behavior had diminished, and he continued to study while drinking his milk and did not play with the carton afterwards. Robbie's intervention was based on the principle of positive reinforcement.

This chapter examines the definition and nature of positive reinforcement, describes methods for identifying potential reinforcers and assessing their effects, outlines experimental control techniques for verifying whether a positive reinforcement contingency is responsible for increased responding, and offers guidelines for using positive reinforcement effectively.

POSITIVE REINFORCEMENT DEFINED

The principle of reinforcement is deceptively simple. "The basic operant functional relation for reinforcement is the following: When a type of behavior (R) is followed by reinforcement (S^R) there will be an increased future frequency of that type of behavior" (Michael, 2004, p. 30).[1] However, as Michael pointed out, three qualifications must be considered regarding the conditions under which the effects of reinforcement will occur: (a) the timing between the end of a given response and the onset of the stimulus change (i.e., the presentation of the reinforcer), (b) the relationship between stimulus conditions present when the response was emitted, and (c) the role of motivation. In this section we examine these qualifications and several other concepts requisite to acquiring a full understanding of reinforcement.

Operation and Defining Effect of Positive Reinforcement

Positive reinforcement occurs when a response is followed immediately by the presentation of a stimulus change that increases the future occurrence of similar responses. Figure 11.2 illustrates the two-term contingency—a response followed closely in time by the presentation of a stimulus—and the increased future responding that defines positive reinforcement. This two-term contingency is the fundamental building block for the selection of all operant behavior.

The stimulus that is presented as a consequence, and that is responsible for the subsequent increase in responding, is called a **positive reinforcer**, or, more simply, a reinforcer. Teacher attention in the form of positive praise was the reinforcer that increased Robbie's study behavior. Cold water flowing into a cup and the sight of a colorful bird are the reinforcers for the two behaviors shown in Figure 11.2.

It is important to remember that a reinforcer does not (and cannot) affect the response that it follows. When making this point, Skinner (1953) reminded us that reinforcement affects the operant response class.

> It is not correct to say the operant reinforcement "strengthens the response which precedes it." The response has already occurred and cannot be changed. What is changed is the future probability of responses in the same class. It is the operant as a class of behavior, rather than the response as a particular instance, which is conditioned. (p. 87)

Skinner (1938, 1966) used rate of responding as the fundamental datum for his research on reinforcement. To strengthen an operant is to make it occur more often.[2] However, rate is not the only dimension of behavior selected, shaped, and maintained by reinforcement. Reinforcement can also alter the duration, latency, interresponse time, magnitude, and/or topography of behavior. For example, if reinforcement follows only responses that fall within a range of magnitude—that is, above a minimum force but below a maximum force—responses within that range will occur more frequently. Reinforcement contingent on responses meeting multiple criteria will alter the response class meeting those criteria (e.g., responses by a golfer practicing 10-foot putts must fall within a narrow range of force and form to be successful).

Immediacy of Reinforcement

The direct effects of reinforcement involve "temporal relations between behavior and its consequences that are on the order of a few seconds" (Michael, 2004, p. 161). Although research with nonhuman animals suggests that as much as 30 seconds can elapse without critical loss of effect (e.g., Bradley & Poling,

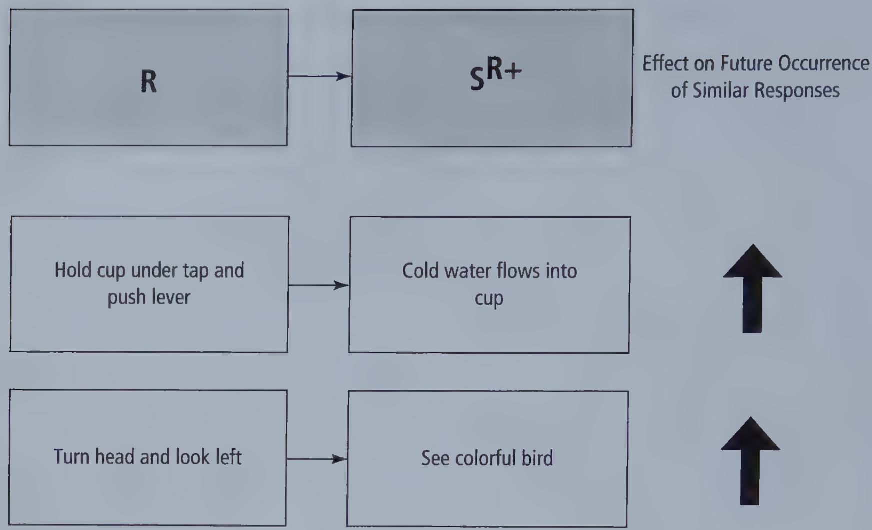


Figure 11.2 Two-term contingency illustrating positive reinforcement: A response (R) is followed closely in time by a stimulus change (S^{R+}) that results in an increased occurrence of similar responses in the future. Used with the permission of Skinner, B. F. (1953). Science and human behavior. New York: Macmillan.

2010; Byrne, LeSage, & Poling, 1997; Critchfield & Lattal, 1993; Lattal & Gleason, 1990; Wilkenfeld, Nickel, Blakely, & Poling, 1992), a response-to-reinforcement delay of as little as 1 second will be less effective than a reinforcer delivered immediately. This is because behaviors other than the target behavior can occur during the delay; the behavior temporally closest to the presentation of the reinforcer will be strengthened by its presentation. As Sidman (1960) described, "If the reinforcer does not immediately follow the response that was required for its production, then it will follow some other behavior. Its major effect will then be upon the behavior that bears, adventitiously to be sure, the closest prior temporal relationship to the reinforcement" (p. 371).

Malott and Shane (2014) discussed the importance of a brief response-to-reinforcement delay, as follows:

> [I]f you're trying to reinforce a response, don't push that 60-second limit, push the other end—the 0-second end. The direct effect of reinforcement drops off quickly as you increase the delay, even to 3 or 4 seconds. And even a 1-second delay may reinforce the wrong behavior. If you ask a young child to look at you and deliver the reinforcer 1 second after the response, you're liable to reinforce looking in the wrong direction. So, one problem with delayed reinforcement is that it reinforces the wrong response—the one that occurred just before the delivery of the reinforcer. (p. 4)

A common misconception is that delayed consequences can reinforce behavior, even when the consequences occur days, weeks, or even years after the responses occurred. "When human behavior is apparently affected by long-delayed consequences, the change is accomplished by virtue of the human's complex social and verbal history, and should not be thought of as an instance of the simple strengthening of behavior by reinforcement" (Michael, 2004, p. 36).

For example, suppose that a piano student practiced dutifully every day for several months in preparation for a statewide competition, at which she received a first-place award for her solo piano performance. Although some might believe that the award reinforced her persistent daily practice, they would be mistaken. Delayed consequences do not reinforce behavior directly. Delayed consequences can, when combined with language, *influence* future behavior through instructional control and rule following. A *rule* is a verbal description of a behavioral contingency (e.g., "Turnip seeds planted by August 15 will yield a crop before a killing freeze."). Learning to follow rules is one way that a person's behavior can come under the control of consequences that are too delayed to influence behavior directly. A statement by the piano teacher such as "If you practice your assignments every day for one hour between now and the competition, you could win first place" could have functioned as a rule that influenced the piano student's daily practice. The student's daily practice was evidence of rule-governed behavior if daily practice occurred because of her teacher's rule.[3] In other words, **rule-governed behavior** is behavior controlled by a rule (i.e., a verbal statement of an antecedent-behavior-consequence contingency) that enables human behavior to come under the indirect control of temporarily remote or improbable but potentially significant consequences. The following conditions provide strong indicators that behavior is the result of instructional control or is rule-governed behavior rather than a direct effect of reinforcement (Malott, 1988; Michael, 2004).

- No immediate consequence for the behavior is apparent.
- The response–consequence delay is greater than 30 seconds.
- Behavior changes without reinforcement.
- A large increase in the occurrence of the behavior occurs following one instance of reinforcement.
- No consequence for the behavior exists, including no automatic reinforcement, but the rule exists.

Reinforcement Is Not a Circular Concept

Another common misconception is that reinforcement is the product of circular reasoning and therefore contributes nothing to our understanding of behavior. Not so. Circular reasoning is a form of faulty logic in which the name used to describe an observed effect is mistaken as the cause for the effect. This confusion of cause and effect is circular because the observed effect is the sole basis for identifying the presumed cause.

In circular reasoning, the suspected cause is not independent of its effect—they are one and the same.

Here is an example of circular reasoning that occurs often in education. A student's persistent difficulties learning to read (effect) leads to a formal diagnosis of a learning disability, which is then offered as an explanation for the student's struggles with reading: "Yoshi's reading problem is due to his learning disability." How do you know Yoshi has a learning disability? Because he hasn't learned to read. Why hasn't Yoshi learned to read? Because his learning disability has prevented him from learning to read. And around and around it goes.

Similarly, it would be circular reasoning if we said that teacher attention increased Robbie's study behavior *because* it is a reinforcer. Instead, it is correct to say that because Robbie's study behavior increased when (and only when) it was followed immediately by teacher attention, teacher attention is a reinforcer. The difference is more than the direction of the relation, or some semantic sleight-of-hand. In circular reasoning, the suspected cause is not manipulated as an independent variable to see whether it affects the behavior. In circular reasoning, such experimental manipulation is impossible because the cause and effect are the same. Yoshi's learning disability cannot be manipulated as an independent variable because, as we used the concept in this example, it is nothing more than another name for the dependent variable (effect).

Reinforcement is not a circular concept because the two components of the response–consequence relation can be separated, allowing the delivery of a consequence to be manipulated to determine whether it increases the occurrence of the behavior it follows. Epstein (1982) described it as follows:

> If we can show that a response increases in frequency because (and only because) it is followed by a particular stimulus, we call that stimulus a *reinforcer* and its presentation, *reinforcement*. Note the lack of circularity. *Reinforcement* is a term we invoke when we observe certain relations between events in the world. . . . [However,] if we say, for example, that a particular stimulus strengthens a response behavior *because* it is a reinforcer, we are using the term *reinforcer* in a circular fashion. It is *because* it strengthens behavior that we call the stimulus a *reinforcer*. (p. 4)

Epstein (1982) went on to explain the difference between using an empirically demonstrated principle such as reinforcement in a theoretical account of behavior and using a circular argument.

> In some of his writings, Skinner speculates that certain behavior (for example, verbal behavior) has come about through reinforcement. He may suggest, for example, that certain behavior is strong *because* it was reinforced. This use of the concept is not circular, only speculative or interpretive. Using the language of reinforcement in this way is reasonable when you have accumulated a large data base. . . . When Skinner attributes some everyday behavior to past reinforcers, he is making a plausible guess based on a large data base and principles of behavior established under controlled conditions. (p. 4)

Used properly, *reinforcement* describes an empirically demonstrated (or speculative, in a theoretical or conceptual analysis) functional relation between a stimulus change (consequence) immediately following a response and an increase in the future occurrence of similar responses. Table 11.1 shows restrictions and examples of appropriate use of the terms *reinforcer, reinforcing, reinforcement*, and *to reinforce*, as suggested by Catania (2013). Box 11.1 describes four mistakes commonly made when speaking and writing about reinforcement.

Relationship Between Reinforcement and Antecedent Stimulus Conditions

Reinforcement does more than increase the future occurrence of behavior; it also changes the function of stimuli that immediately precede the reinforced behavior. By virtue of being

TABLE 11.1 The Vocabulary of Reinforcement*

Term	Restrictions	Examples
reinforcer (noun)	A stimulus	Food pellets were used as reinforcers for the rat's lever presses.
reinforcing (adjective)	A property of a stimulus	The reinforcing stimulus was produced more often than the other, nonreinforcing stimuli.
reinforcement (noun)	As an operation, the delivery of consequences when a response occurs	The fixed-ratio schedule of reinforcement delivered food after every 10th key peck.
	As a process, the increase in responding that results from the reinforcement	The experiment with monkeys demonstrated reinforcement produced by social consequences.
to reinforce (verb)	As an operation, to deliver consequences when a response occurs; responses are reinforced and not organisms	When a period of free play was used to reinforce the child's completion of school work, the child's grades improved.
	As a process, to increase responding through the reinforcement operation	The experiment was designed to find out whether gold stars would reinforce cooperative play among first-graders.

*This vocabulary is appropriate if and only if three conditions exist: (1) A response produces consequences; (2) that type of response occurs more often when it produces those consequences than when it does not produce them; and (3) the increased responding occurs *because* the response has those consequences. A parallel vocabulary is appropriate to punishment (including *punisher* as a stimulus and *punish* as a verb), with the difference being that a punishing consequence reduces the likelihood of future responding.

Based on *Learning, Interim* (5th ed.) by A. C. Catania, 2013, p. 66. Cornwall-on-Hudson, NY: Sloan Publishing.

BOX 11.1

Common Mistakes in Talking and Writing About Reinforcement

A standard set of technical terms is prerequisite to the meaningful description of any scientific activity. Effectively communicating the design, implementation, and outcome of an applied behavior analysis depends on the accurate use of the discipline's technical language. The language of reinforcement includes some of the most important elements of the behavior analyst's vocabulary.

In this box we identify four mistakes made frequently by students of applied behavior analysis when describing reinforcement-based interventions. Perhaps the most common mistake—confusing negative reinforcement with punishment—is not discussed here. That terminology error was introduced in Chapter 2 and receives additional attention in Chapter 12.

Reinforcing the Person

Although it is proper to speak of presenting a *reinforcer* to a learner (e.g., "The teacher delivered a token to Bobby each time he asked a question"), statements such as "The teacher reinforced Bobby when he asked a question" and "Chloe was reinforced with praise each time she spelled a word correctly" are incorrect. *Behaviors* are reinforced, not people. Bobby's teacher reinforced question asking, not Bobby. Of course, reinforcement acts on and affects the overall person, in that it strengthens behaviors within the person's repertoire. However, the procedural focus and the primary effect of reinforcement are on the behaviors that it follows.

Practice as Reinforcement for a Skill

Educators will sometimes say that students should practice a skill because "practicing reinforces the skill." The phrase poses no problem if the speaker is describing a common outcome of practice with the everyday language connotation of *reinforce*, as in "to make something stronger" (e.g., to reinforce concrete by embedding steel rods in it). Well-designed drill and practice on a skill usually yields stronger performance in the form of better retention, reduced latency, higher response rates, and/or increased endurance (e.g., Johnson & Layng, 1994; Swanson & Sachse-Lee, 2000). Unfortunately, a phrase such as "practicing reinforces the skill" is often misused and misinterpreted as technical usage of the language of operant conditioning.

Although a skill that has been practiced is often stronger as a result of the practice, the practice itself could not be a *reinforcer* for the behavior practiced. Practice refers to the form and manner in which the target skill is emitted (e.g., answering as many math problems as you can in 1 minute). Practicing is a behavior that could be reinforced with various consequences, such as a preferred activity (e.g., "Practice solving these math problems; then you can have 10 minutes of free time"). Depending on a learner's history and preferences, the opportunity to practice a certain skill may function as a reinforcer for practicing another skill (e.g., "Finish your math problems; then you'll get to do 10 minutes of repeated reading practice").

Artificial Reinforcement

A distinction between natural and artificial reinforcers is made sometimes, as in this statement: "As the students' success rates improved, we gradually stopped using artificial reinforcers, such as stickers and trinkets, and increased the use of natural reinforcers." Some authors have suggested that applications of the principles of behavior result in "artificial control" (e.g., Smith, 1992). A behavior–consequence contingency may be effective or ineffective as reinforcement, but none of its elements (the behavior, the consequence, or the resultant behavior change) is, or can be, artificial.

The reinforcement contingencies and stimuli used as reinforcers in any behavior change program are always contrived—otherwise, there would be no need for the program—but they are never artificial (Skinner, 1982). The meaningful distinction when talking about reinforcement contingencies is not between the natural and the artificial, but between contingencies that already exist in a given setting prior to a behavior change program and contingencies that are contrived as part of the program (Kimball & Heward, 1993). Although the ultimate effectiveness of a behavior change program may depend on shifting control from contrived to naturally occurring contingencies, there is no such thing as artificial reinforcement.

Reinforcement and Feedback as Synonyms

Some speakers and writers mistakenly use *reinforcement* and *feedback* interchangeably. The two terms refer to different operations and outcomes, though some of each term encompasses parts of the other term's meaning. *Feedback* is information a person receives about a particular aspect of his or her behavior following its completion (e.g., "Very good, Kathy. Two quarters equal 50 cents."). Feedback is most often provided in the form of verbal descriptions of performance, but it can also be provided by other means, such as vibration or lights (e.g., Greene, Bailey, & Barber, 1981). Because feedback is a consequence that often results in the increased future occurrence of behavior, it sometimes leads to the faulty assumption that reinforcement must involve feedback or that *reinforcement* is just a behaviorist's term for *feedback*.

Reinforcement always increases the future frequency of responding. Feedback may result in (a) an increase in the future occurrence of the student's performance as a reinforcement effect and/or as a prompt or instruction on how to respond next time (e.g. "Your handwriting is improving, Jason, but don't forget to cross your T's") and/or (b) a reduction in the occurrence of some aspect of the learner's performance as a function of punishment or instruction (e.g., "You dropped your elbow on that pitch. Don't do that."). Feedback may have multiple effects, increasing one aspect of performance and decreasing another. Feedback may also have no effect on future responding whatsoever.

Reinforcement is defined functionally by its effect on future responding; feedback is defined by its formal characteristics (information about some aspect of performance). The operation of either concept is neither necessary nor sufficient for the other. That is, reinforcement may occur in the absence of feedback, and feedback may occur without a reinforcement effect.

Sometimes Commonsense Language Is Better

The technical language of behavior analysis is complex, and mastering it is no simple matter. Beginning students of behavior analysis are not the only ones who commit terminology errors. Well-trained practitioners, established researchers, and experienced authors also make mistakes now and then when speaking and writing about behavior analysis. Using behavioral concepts and principles—such as positive reinforcement—to confidently explain complex situations involving multiple processes and uncontrolled and unknown variables is a mistake that catches the most attentive and conscientious of behavior analysts at times.

Instead of invoking the terminology and concepts of reinforcement to explain the influence of temporally distant consequences on behavior, it is probably wiser to follow Jack Michael's (2004) advice and simply use everyday descriptive language and commonsense relations.

> Incorrectly used technical language is worse than common sense language because it suggests that the situation is well understood, and it may displace serious attempts at further analysis. Until we are able to provide an accurate analysis of the various processes relevant to indirect effects [of reinforcement], we are better off using ordinary descriptive language. Thus, say "the successful grant application is *likely to encourage* future efforts in the same direction," but don't say it as though you had the science of behavior behind you. Stop referring to successful settlements of a labor dispute as reinforcement for striking, and successful election of a political candidate as reinforcement for political activity. . . . Don't talk about good grades as reinforcement for effective study behavior, although they are no doubt responsible for maintaining it in some cases. Just say that they're responsible for maintaining it. Restraint of this sort will deprive some of us of an opportunity to (incorrectly) display our technical knowledge, but so much the better. (p. 165, emphasis in original)

temporally paired with the response–reinforcer contingency, antecedent events acquire the ability to evoke (make more likely) instances of the reinforced response class. As introduced in Chapter 2, a *discriminative stimulus* (S^D, pronounced "ess-dee") is an antecedent stimulus correlated with the availability of reinforcement for a particular response class. Responding in the presence of the S^D produces reinforcement; responding in the absence of the S^D (a condition called *stimulus delta* (S^Δ, pronounced "ess-delta") does not. As a result of this history of reinforcement, a person learns to make more responses in the presence of the S^D than in its absence. The behavior is then considered to be under *stimulus control* (see Chapter 17).

With the addition of the S^D, the two-term contingency for reinforcement becomes the three-term contingency of the *discriminated operant*. Figure 11.3 shows examples of three-term contingencies for positive reinforcement. Assuming that cold water is currently reinforcing and the person has a history of obtaining cold water only under blue taps, he is more likely to hold his cup under the blue tap on the cooler (than, say, a red tap). Similarly, assuming that seeing a colorful bird is currently reinforcing and a person has a history of seeing birds more often when looking toward chirping sounds (than, say, other sounds or silence), turning one's head and looking to the left will occur more often when chirping is heard to the left.

The Role of Motivation

The phrase *assuming that cold water is currently reinforcing* in the previous paragraph holds another key to understanding reinforcement. Although reinforcement is commonly thought of as a way to motivate people—and it can be—the momentary effectiveness of any stimulus change as reinforcement depends on an existing level of motivation with respect to the stimulus change in question. As introduced in Chapter 2, *motivating operations* alter the current effectiveness of stimulus changes as reinforcement.

> Motivating operations (MOs) are environmental variables that have two effects on behavior: (1) They alter the operant reinforcing effectiveness of some specific stimuli, objects, or events (the value-altering effect); and (2) They alter the momentary frequency of all behavior that has been reinforced by those stimuli, objects, or events (the behavior-altering effect). The value-altering effect, like response-reinforcement delay, is relevant to the effectiveness of the reinforcer at the time of conditioning, and stating that the consequence is a form of reinforcement implies that a relevant MO is in effect and at sufficient strength. (Michael 2004, p. 31)

In other words, for a stimulus change to "work" as reinforcement at any given time, the learner must already *want* it. This is a critical qualification in terms of the environmental conditions under which the effects of reinforcement will be seen. Michael (2004) explained this qualification as follows:

> The behavior-altering effect is relevant to the increased future frequency of the reinforced behavior, and must be added as a third qualification to the operant reinforcement relation: In a given stimulus situation (S) when a type of behavior (R) is followed immediately by reinforcement (S^R) there will be an increase in the future frequency of that type of behavior in the same or similar stimulus conditions, *but the increased frequency will only be seen when the MO relevant to the reinforcement that was used is again in effect.* (p. 31, emphasis in original)

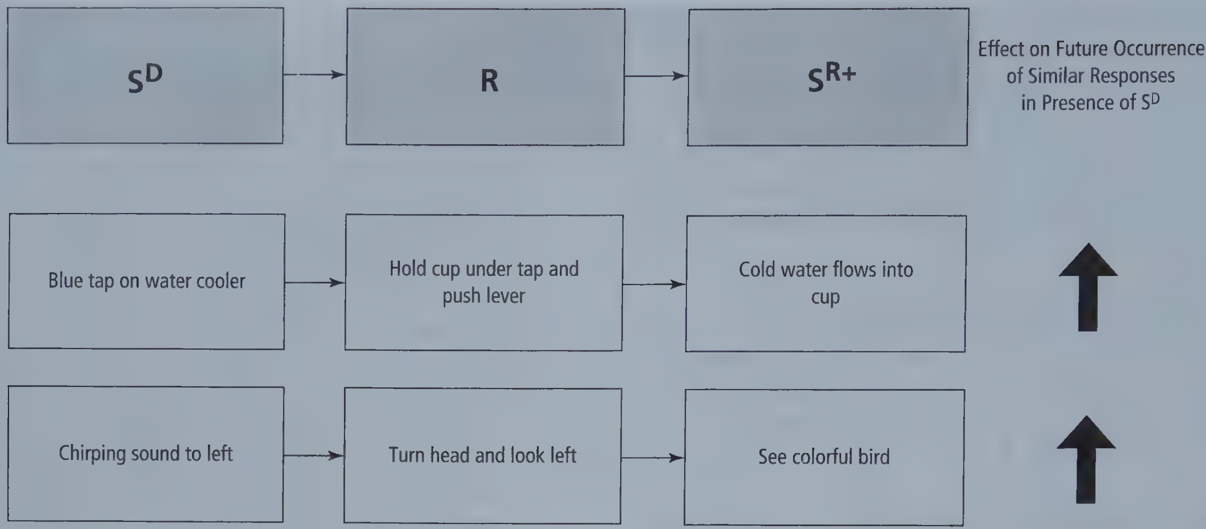


Figure 11.3 Three-term contingency illustrating positive reinforcement of a discriminated operant: A response (R) emitted in the presence of a discriminative stimulus (S^D) is followed closely in time by a stimulus change (S^{R+}) and results in an increased occurrence of similar responses in the future when the S^D is present. A discriminated operant is the product of a conditioning history in which responses in the presence of the S^D produced reinforcement, and responses in the absence of the S^D (a condition called stimulus delta [S^{Δ}]) were not reinforced (or resulted in a reduced amount or quality of reinforcement than in the S^D condition).

Motivating operations take two forms. An MO that increases the current effectiveness of a reinforcer is called an *establishing operation (EO)* (e.g., food deprivation makes food more effective as a reinforcer); an MO that decreases the current effectiveness of a reinforcer is an *abolishing operation (AO)* (e.g., food ingestion reduces the effectiveness of food as a reinforcer).[4]

Adding the *establishing operation (EO)* to a discriminated operant results in a four-term contingency, as shown in Figure 11.4. Spending several hours in a hot and stuffy room without water is an EO that (a) makes water more effective as a reinforcer and (b) increases the momentary frequency of all behaviors that have produced water in the past. Similarly, a park ranger stating prior to a hike that any hiker who describes the coloring of the bird that makes a certain chirping sound will receive a $5 token for the gift shop is an EO that will (a) make seeing a bird that makes the chirping sound effective as reinforcement and (b) increase the frequency of all behaviors (e.g., turning one's head to look around, stepping quietly) that have produced similar consequences (in this case, seeing the source of sounds) in the past.

In simpler terms, establishing operations (EOs) determine what an individual *wants* at any particular moment. EOs are dynamic, always changing. Reinforcer effectiveness

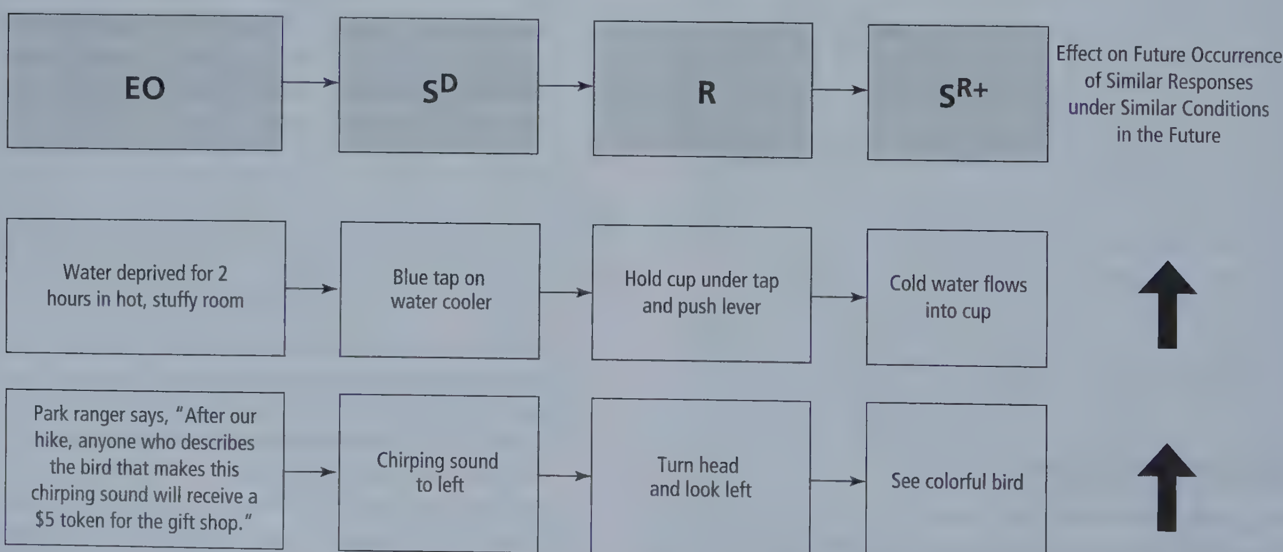


Figure 11.4 Four-term contingency illustrating positive reinforcement of a discriminated operant made current by a motivating operation: An establishing operation (EO) increases the momentary effectiveness of a stimulus change as a reinforcer, which in turn makes the S^D more likely to evoke behavior that has been reinforced by that stimulus change in the past.

(the want) goes up with increasing levels of deprivation and goes down with levels of satiation. Vollmer and Iwata (1991) demonstrated how the reinforcing effectiveness of three classes of stimuli—food, music, and social attention—varied under conditions of deprivation and satiation. Participants were five adults with developmental disabilities, and the dependent variable was the number of responses per minute on two motor tasks—pressing a switch or picking up small blocks from a container and putting them through the hole in the top of another container. All sessions lasted 10 minutes and began with the experimenter saying, "Do this, [participant's name]," and modeling the response. During baseline, participants' responses received no programmed consequences. During the deprivation and satiation conditions, responses were followed by presentation of food, music, or social attention. Initially each response was followed by the programmed consequence; this gradually shifted to every third, fifth, or tenth response being followed by the consequence.

Different procedures were used to create deprivation and satiation conditions for each stimulus class. With food, for example, baseline and deprivation condition sessions were conducted 30 minutes prior to a participant's scheduled lunchtime; sessions during the satiation condition were conducted within 15 minutes after the participant had eaten lunch. For social attention, baseline and deprivation condition sessions were conducted immediately following a 15-minute period in which the participant had either been alone or been observed to have had no social interaction with another person. Immediately prior to each session in the satiation condition, the experimenter provided continuous social interaction (e.g., played a simple game, conversation) with the participant for 15 minutes.

All five participants responded at higher rates under the deprivation conditions than during the satiation conditions. Figure 11.5 shows the effects of deprivation and satiation of social attention on the effectiveness of social attention as a reinforcer for two of the study's participants, Donny and Sam. Other researchers have reported similar findings concerning the effects of deprivation and satiation of various stimuli and events as motivating operations that affect the relative effectiveness of reinforcement (e.g., Hanley, Iwata, & Roscoe, 2006; Klatt, Sherman, & Sheldon, 2000; North & Iwata, 2005; Zhou, Iwata, & Shore, 2002).

Automaticity of Reinforcement

> *A reinforcing connection need not be obvious to the individual reinforced.*
>
> —B. F. Skinner (1953, p. 75)

The fact that a person does not have to understand or verbalize the relation between his or her actions and a reinforcing consequence, or for that matter even be aware that a consequence has occurred, for reinforcement to occur is known as the *automaticity of reinforcement*. Skinner (1983) provided an interesting example of automaticity in the third and final volume of his autobiography, *A Matter of Consequences*. He described an incident that took place at a meeting of distinguished scholars who had been invited to discuss the role of intention in political activity. At one point during the meeting, the psychologist Erich Fromm began to argue that "people were not pigeons," perhaps

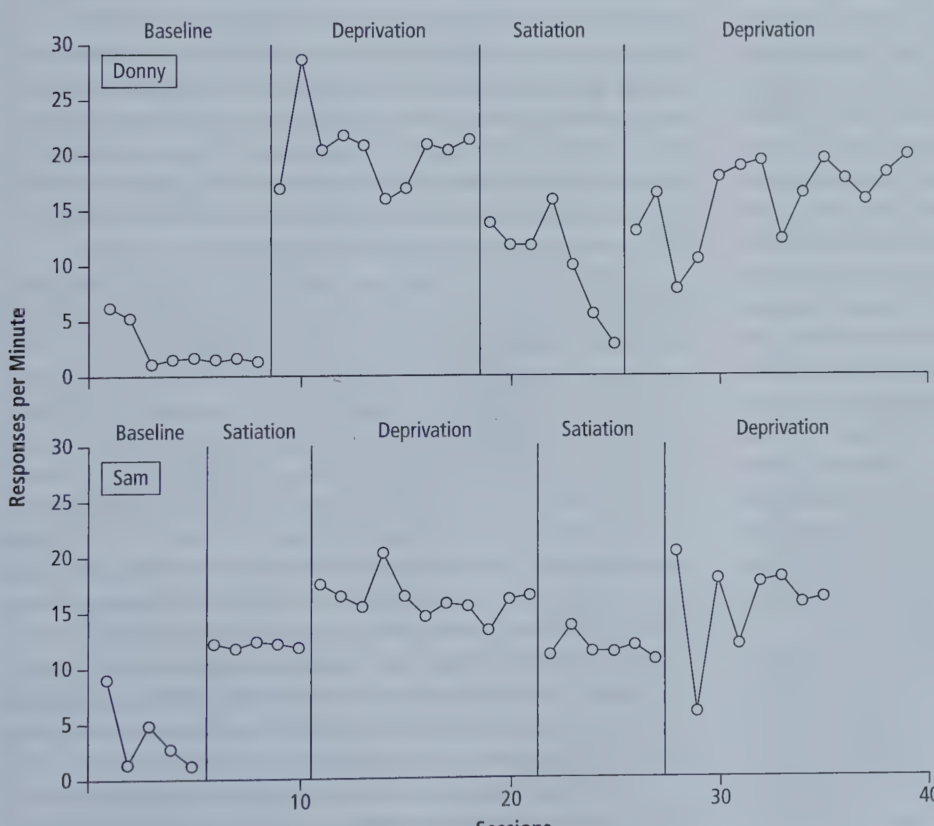


Figure 11.5 Responses per minute by two students during baseline and when social attention was used as reinforcement under deprivation and satiation conditions for social attention.

From "Establishing Operations and Reinforcement Effects," by T. R. Vollmer and B. A. Iwata, 1997, *Journal of Applied Behavior Analysis*, *24*, p. 288. Copyright 1991 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

implying that an operant analysis based on positive reinforcement could not explain human behavior, which Fromm believed is the product of thought and free will. Skinner recounted what happened next:

> I decided that something had to be done. On a scrap of paper I wrote, "Watch Fromm's left hand. I am going to shape [reinforce by successive approximations] a chopping motion" and passed it down the table to Halleck [a member of the group]. Fromm was sitting directly across the table and speaking mainly to me. I turned my chair slightly so that I could see him out of the corner of my eye. He gesticulated a great deal as he talked, and whenever his left hand came up, I looked straight at him. If he brought the hand down, I nodded and smiled. Within five minutes he was chopping the air so vigorously that his wristwatch kept slipping out over his hand. (pp. 150–151, words in brackets added)

Arbitrariness of the Behavior Selected

> *So far as the organism is concerned, the only important property of the contingency is temporal.*
>
> —B. F. Skinner (1953, p. 85)

Reinforcement occurs independent of a logical or adaptive connection between behavior and a reinforcing consequence. In other words, reinforcement strengthens any behavior that immediately precedes it. This arbitrary nature of the behavior selected is critical to understanding reinforcement. All other relations (e.g., what's logical, desirable, useful, appropriate) must compete with the temporal relation between behavior and consequence. "To say that reinforcement is contingent upon a response may mean nothing more than that it followed the response . . . conditioning takes place presumably because of the temporal relation only, expressed in terms of the order and proximity of response and reinforcement" (Skinner, 1948, p. 168).

Skinner (1948) demonstrated the arbitrary nature of the behaviors selected by reinforcement in one of his most famous experimental papers, "Superstition in the Pigeon." He gave pigeons a small amount of food every 15 seconds, "with no reference whatsoever to the bird's behavior" (p. 168). The fact that reinforcement will strengthen whatever behavior it immediately follows was soon evident. Six of the eight birds developed idiosyncratic behaviors "so clearly defined, that two observers could agree perfectly in counting instances" (p. 168). One bird walked counterclockwise around the cage; another repeatedly thrust its head into one of the upper corners of the cage. Two birds acquired a "pendulum motion of the head and body, in which the head was extended forward and swung from right to left with a sharp movement followed by a somewhat slower return" (p. 169). The pigeons had exhibited none of those behaviors at "any noticeable strength" during adaptation to the cage or before the food was periodically presented.

Whatever behavior the pigeons happened to be doing when the food hopper appeared tended to be repeated, which made it more likely to be occurring when food appeared the next time. Reinforcement was not contingent (in the sense of, dependent) on the behavior; it was only a coincidence that reinforcement sometimes followed the behavior. Such accidentally reinforced behavior is called "superstitious" because it has no influence on whether reinforcement follows. Humans engage in many superstitious behaviors. Sports provide countless examples: A basketball player tugs on his shorts before shooting a foul shot, a golfer carries his lucky ball marker, a batter goes through the same sequence of adjusting his wristbands before each pitch, a college football fan wears a goofy-looking necklace made of inedible nuts to bring good luck to his team.[5]

The importance of understanding the arbitrariness of reinforcement goes far beyond providing a possible explanation for the development of harmless superstitious and idiosyncratic behaviors. The arbitrary nature of selection by reinforcement may explain the acquisition and maintenance of many maladaptive and challenging behaviors. For example, a caregiver's well-meaning social attention provided in an attempt to console or divert a person who is hurting himself may help shape and maintain the very behavior the caregiver is trying to prevent or eliminate. Kahng, Iwata, Thompson, and Hanley (2000) documented with a functional analysis that social reinforcement maintained the self-injurious behavior (SIB) and aggression of three adults with developmental disabilities. Kahng and colleagues' data support the hypothesis that aberrant behaviors may be selected and maintained by social attention because of the arbitrariness of reinforcement.

Automatic Reinforcement

Kennedy (1994) noted that applied behavior analysts use two meanings to define the term *automatic reinforcement*. In the first instance, automatic reinforcement is determined by the absence of social mediation (Vollmer, 1994, 2006). In this context, **automatic reinforcement** refers to the behavior–stimulus change relation that occurs without the presentation of consequences by other people (Vaughan & Michael, 1982; Vollmer, 1994, 2006). In short, automatic reinforcement occurs independent of social mediation of others. Response products that function as automatic reinforcement are often in the form of a naturally produced sensory consequence that "sounds good, looks good, tastes good, smells good, feels good to the touch, or the movement itself is good" (Rincover, 1981, p. 1). Scratching an insect bite to relieve or attenuate the itch is a common example of automatic reinforcement.

In the second instance, *automatic reinforcement* is assumed when a behavior persists in the absence of any known reinforcer (Hagopian, Rooker, & Yenokyan, 2018). In these "default" situations, automatic reinforcement is assumed to be the controlling variable (Fisher, Lindauer, Alterson, & Thompson, 1998; Ringdahl, Vollmer, Marcus, & Roane, 1997; Roscoe, Iwata, & Goh, 1998). Persistent, seemingly nonpurposeful, repetitive, self-stimulatory behaviors (e.g., flipping fingers, head rolling, body rocking, toe walking, hair pulling, fondling body parts) may produce sensory stimuli that function as automatic reinforcement. Such "self-stimulation" is thought to be a factor maintaining self-injurious behavior (Iwata, Dorsey, Slifer, Bauman, & Richman, 1994), stereotypic repetitive movements, hair pulling (Rapp, Miltenberger, Galensky, Ellingson, & Long, 1999), nail biting, chewing on the mouth or lips, and twirling objects or fondling jewelry (Miltenberger, Fuqua, & Woods, 1998).

Determining that a behavior may be maintained by automatic reinforcement, and when possible isolating or substituting the source of that reinforcement (see Figure 8.5; Shore,

Iwata, DeLeon, Kahng, & Smith, 1997), has important implications for designing interventions to either capitalize on the automatic reinforcing nature of the behavior or counteract it. As Vollmer (1994) stated ". . . one clear goal of research should be to develop procedures designed to identify specific sources of automatic reinforcement. Effective assessment might allow for the development of treatments that involve reversing or eliminating the contingencies maintaining a behavior" (p. 191). Assuming that such assessment provides important clues for intervention, Vollmer further proposed five interventions to reduce or eliminate behaviors maintained by automatic reinforcement.

1. Eliminate or minimize medically based establishing operations. For example, a child whacking his ear or scratching his arm repeatedly might be treated by an antibiotic or medicated lotion to reduce or eliminate aversive stimulation produced by an earache or allergy-based itch that established whacking or scratching responses as a negative reinforcement.
2. Try a sensory extinction procedure, thus eliminating, or reducing, sources of visual, auditory, or proprioceptive stimulation. Wearing gloves, shielded clothes, helmets, and so forth serve as examples.
3. Implement differential reinforcement of alternative behavior (DRA) interventions, thus providing a "choice" with an alternative source of reinforcement contingent on appropriate behavior. A course of action would provide a different source of visual, auditory, or proprioceptive stimulation that would compete with the automatically reinforced behavior. Teaching the person to turn on an electric vibrator, television, or musical carousel to obtain alternative stimulation serve as examples.
4. Consider punishment to override current maintaining contingencies, and thus capitalize on the fact that (a) punishment does not require the knowledge of the specific controlling variables, and (b) exigent circumstances do not allow for prolonged detective work to determine the existing reinforcement contingencies.
5. Implement a "package" program. "It is unlikely that any given intervention derived from establishing operations, sensory extinction, differential reinforcement, or punishment would be used in isolation from other procedures. Package treatment approaches are common both in research and in clinical settings. In the case of automatic reinforcement, it is possible that a package approach would be highly suitable". (Vollmer, 1994, p. 203)

Each of these recommendations has advantages and disadvantages, and none is a "silver bullet" that will instantly address automatic reinforcement's effect on behavior. As Vollmer (1994) notes, "When reinforcers maintaining a behavior are not within the control of the therapist or experimenter (i.e., the social environment), behavior is especially difficult to assess and treat" (p. 188). Still, these five recommendations provide a direction for analysts to determine their efficacy in individual instances.

Responses may produce unconditioned or conditioned automatic reinforcement. Sundberg, Michael, Partington, and Sundberg (1996) described a two-stage conditioning history that may account for conditioned automatic reinforcement.

> For example, a person may persist in singing or humming a song while coming home from a movie despite no obvious direct reinforcement for singing. In order for this behavior to occur as automatically reinforced behavior, a special two-stage conditioning history is necessary. In stage one, some stimulus (e.g., a song) must be paired with an existing form of conditioned or unconditioned reinforcement (e.g., an enjoyable movie, popcorn, relaxation). As a result, the new stimulus can become a form of conditioned reinforcement (e.g., hearing the song may now be a new form of conditioned reinforcement). In stage two, the emission of a response (for whatever reason) produces a response product (i.e., the auditory stimuli produced by singing the song) that has topographical similarity to that previously neutral stimulus (e.g., the song), and may now have self-strengthening properties. (pp. 22–23)

Several theorists have suggested that automatic reinforcement may help to explain the extensive babbling of infants and how babbling shifts naturally, without apparent intervention from others, from undifferentiated vocalizations to the speech sounds of their native language (e.g., Bijou & Baer, 1965; Mowrer, 1950; Skinner, 1957; Staats & Staats, 1963; Vaughan & Michael, 1982). Caregivers frequently talk and sing while holding, feeding, and bathing a baby. As a result of repeated pairing with various reinforcers (e.g., food, warmth), the caregiver's voice may become a conditioned reinforcer for the baby. The baby's babbling is automatically reinforced when it produces sounds that match or closely approximate the caregiver's. At that point, "The young child alone in the nursery may automatically reinforce his own exploratory vocal behavior when he produces sounds that he has heard in the speech of others" (Skinner, 1957, p. 58).

Although the idea that automatic reinforcement is a factor in early language acquisition has been proposed repeatedly, few experimental analyses of the phenomenon have appeared in the literature (e.g., Miguel, Carr, & Michael, 2002; Sundberg, Michael, Partington, & Sundberg, 1996; Yoon & Bennett, 2000). Sundberg and colleagues (1996) reported the first study showing the effects of a stimulus–stimulus pairing procedure on the rate with which children emitted new vocal sounds without direct reinforcement or prompts to respond. Five children, ages 2 to 4 and representing a broad range of language abilities, served as subjects. During the prepairing (baseline) condition, the parents and in-home trainers sat a few feet away from the child and recorded each word or vocal sound emitted by the child as he played with a train set and several toys. Data were collected in consecutive 1-minute intervals. The adults did not interact with the subject during the prepairing baseline. The stimulus–stimulus pairing procedure consisted of a familiar adult approaching the child, emitting a target vocal sound, word, or phrase, and then immediately delivering a stimulus that had been established previously as a form of reinforcement for the child (e.g., tickles, praise, bouncing in a parachute held by adults). This stimulus–stimulus pairing procedure was repeated 15 times per minute for 1 or 2 minutes. The adult used a variety

of pitches and intonations when voicing the target sound, word, or phrase. During the postpairing condition, which began immediately after the stimulus–stimulus pairings, the adult moved away from the child and conditions were the same as during the prepairing condition.

The stimulus–stimulus pairing of a vocal sound, word, or phrase with an established reinforcer was followed by an increased frequency of the targeted word during the postpairing condition for all five children. Figure 11.6 shows the results of a representative sample of one of three pairings conducted with Subject 2, a 4-year-old boy with autism. Subject 2 had a verbal repertoire of more than 200 mands, tacts, and intraverbals, but rarely emitted spontaneous vocalizations or engaged in vocal play.[6] During the prepairing condition, the child did not say the target word and emitted four other vocalizations at a mean rate of 0.5 per minute. The stimulus– stimulus pairing procedure consisted of pairing the word *apple* with tickles approximately 15 times in 60 seconds. Immediately after the pairing, the subject said "apple" 17 times in 4 minutes, a rate of 4.25 responses per minute. In addition, the child said "tickle" four times within the first minute of the postpairing condition. Sundberg and colleagues' results provide evidence that the children's vocal response products may have functioned as automatic conditioned reinforcement after being paired with other forms of reinforcement.

In summing up the uses and limitations of automatic reinforcement as a concept, Vollmer (2006) suggested that:

- Practitioners should recognize that not all reinforcement is planned or socially mediated.
- Some behaviors maintained by automatic reinforcement (e.g., self-stimulation, stereotypy) may not be reduced or eliminated with timeout, planned ignoring, or extinction.
- Affixing the label *automatic reinforcement* to an observed phenomenon too quickly may limit our analysis and effectiveness by precluding further efforts to identify the actual reinforcer-maintaining behavior.
- When socially mediated contingencies are difficult to arrange or simply not available, practitioners might consider automatic reinforcement as a potential aim.

CLASSIFYING REINFORCERS

In this section we review the technical classification of reinforcers by their origin as well as several practical categories by which practitioners and researchers often describe and classify reinforcers by their formal characteristics. Regardless of type or classification, all reinforcers share their most important and defining characteristic: All reinforcers increase the future occurrence of behavior that immediately precedes them.

Classification of Reinforcers by Origin

As introduced in Chapter 2, there are two basic types of reinforcers—a reinforcer that is the product of the evolution of the species (an unconditioned reinforcer) or a reinforcer that is the result of the learning history of the individual (a conditioned reinforcer).

Unconditioned Reinforcers

An **unconditioned reinforcer** is a stimulus change that functions as reinforcement even though the learner has had no particular learning history with it. (The terms *primary reinforcer* and

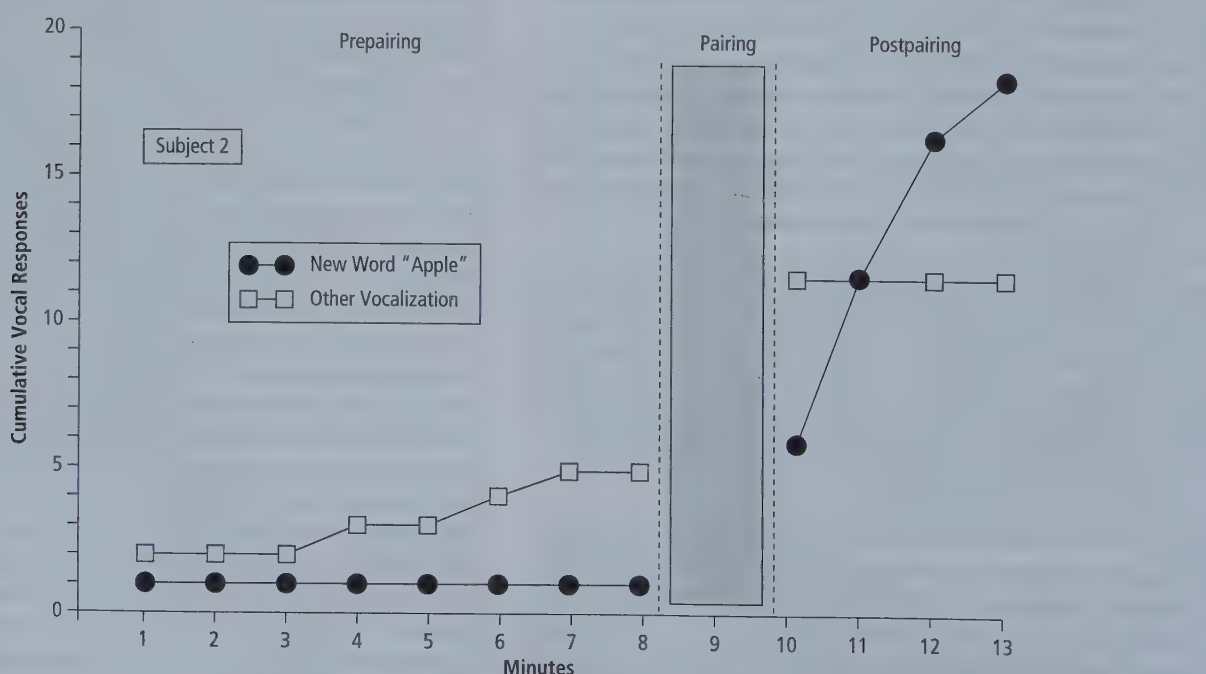


Figure 11.6 Cumulative number of times a 4-year-old child with autism vocalized "apple" before and after "apple" had been paired repeatedly with an established form of reinforcement. Automatic reinforcement may explain the increased frequency of the child's vocalizing "apple" after pairing.

From "The Role of Automatic Reinforcement in Early Language Acquisition," by M. L. Sundberg, J. Michael, J. W. Partington, and C. A. Sundberg, 1996, *The Analysis of Verbal Behavior*, *13*, p. 27. Copyright 1996 by the Association for Behavior Analysis, Inc. Used by permission.

unlearned reinforcer are synonyms for unconditioned reinforcer.) Because unconditioned reinforcers are the product of the evolutionary history of a species (phylogeny), all biologically intact members of a species are more or less susceptible to reinforcement by the same unconditioned reinforcers. For example, food, water, oxygen, warmth, and sexual stimulation are examples of stimuli that do not have to undergo a learning history to function as reinforcers. Food will function as an unconditioned reinforcer for a human deprived of sustenance; water will function as an unconditioned reinforcer for a person deprived of liquid, and so forth.

Human touch may also be an unconditioned reinforcer (Gewirtz & Pelaez-Nogueras, 2000). Pelaez-Nogueras and colleagues (1996) found that infants preferred face-to-face interactions that included touch stimulation. Two conditioning treatments were implemented in alternated counterbalanced order. Under the touch condition, infants' eye-contact responses were followed immediately by adult attention (eye contact), smiling, cooing, and rubbing the infants' legs and feet. Eye-contact responses during the no-touch condition were followed by eye contact, smiles, and coos from the adult, but no touching. All of the babies in the study emitted eye contact for longer durations, smiled and vocalized at higher rates, and spent less time crying and protesting in the contingent condition that included touch. From these results and several related studies, Pelaez-Nogueras and colleagues concluded that "these results suggest that . . . touch stimulation can function as a primary reinforcer for infant behavior" (p. 199).

Conditioned Reinforcers

A **conditioned reinforcer** (sometimes called a *secondary reinforcer* or *learned reinforcer*) is a previously neutral stimulus change that has acquired the capability to function as a reinforcer through stimulus–stimulus pairing with one or more unconditioned reinforcers or conditioned reinforcers. Through repeated pairings, the previously neutral stimulus acquires the reinforcement capability of the reinforcer(s) with which it has been paired.[7] For example, after a tone has been paired repeatedly with food, when food is delivered as a reinforcer, the tone will function as a reinforcer when an EO has made food a currently effective reinforcer.

Neutral stimuli can also become conditioned reinforcers for humans without direct physical pairing with another reinforcer through a pairing process Alessi (1992) called *verbal analog conditioning*.

> For example, a class of preschool children who have been receiving M&M candies for good school work might be shown pieces of cut up yellow construction paper and told, "These pieces of yellow paper are what big kids work for" (Engelmann, 1975, pp. 98–100). Many children in the group immediately refuse M&Ms, and work extra hard, but accept only pieces of yellow paper as their rewards.
>
> We might say that the pieces of yellow paper act as "learned reinforcers." Laboratory research tells us that neutral stimuli become reinforcers only through direct pairing with primary reinforcers (or other "learned reinforcers"). Yellow paper was not paired with any reinforcer and certainly not with the primary (M&Ms) reinforcers. Yellow paper acquired reinforcing properties even more powerful than the primary M&Ms reinforcers, as demonstrated by the children's refusal to accept M&Ms, demanding instead pieces of yellow paper. (For the sake of this example, assume that the children had not been satiated with M&Ms just before the session.) (p. 1368)

It is sometimes thought that the "power" of a conditioned reinforcer is determined by the number of times it has been paired with other reinforcers. However, a statement such as, "The more often the tone is paired with food, the more reinforcing the tone will become" is not completely accurate. Although numerous pairings will increase the likelihood that the tone will function as a conditioned reinforcer in the first place (though a single pairing is sometimes sufficient), the momentary effectiveness of the tone as a reinforcer will be a function of the relevant EO for the reinforcer(s) with which the conditioned reinforcer has been paired. A tone that has been paired only with food will function as an effective reinforcer for a food-deprived learner, but the tone will have little effect as a reinforcer if the learner has just consumed a lot of food, regardless of the number of times it has been paired with food.

A **generalized conditioned reinforcer** is a conditioned reinforcer that as a result of having been paired with many unconditioned and conditioned reinforcers does not depend on a current EO for any particular form of reinforcement for its effectiveness. For example, social attention (proximity, eye contact, praise) is a generalized conditioned reinforcer for many people because it has occurred simultaneously with many reinforcers. The more reinforcers with which a generalized conditioned reinforcer has been paired, the greater the likelihood that it will be effective at any given time. Because it can be exchanged for a nearly limitless variety of backup reinforcers, money is a generalized conditioned reinforcer whose effectiveness is usually independent of current establishing operations.

It is sometimes thought that a conditioned reinforcer is called a generalized conditioned reinforcer because it can function as reinforcement across a wide range of behaviors. This, however, is not so. Any reinforcer is capable of strengthening the future occurrence of any similar behavior that immediately precedes its occurrence. A conditioned reinforcer is called a generalized conditioned reinforcer because it is effective as reinforcement across a wide range of EO conditions. Because of their versatility across EO conditions, generalized conditioned reinforcers offer great advantages for practitioners, who often have limited control of the EOs for particular reinforcers.

Generalized conditioned reinforcers provide the basis for implementing a *token economy*, a reinforcement-based system capable of improving multiple behaviors of multiple participants (Alstot, 2015; Campbell & Anderson, 2011; Hirst, Dozier, & Payne, 2016). In a token economy, participants receive tokens (e.g., points, checkmarks, poker chips) contingent on a variety of target behaviors. Participants accumulate the tokens and exchange them at specific times for their choices from a menu of backup reinforcers (e.g., free time, computer time, snacks). Token economy systems and guidelines for designing and implementing them are described in Chapter 28.

Classification of Reinforcers by Formal Properties

When applied behavior analysts describe reinforcers by their physical properties—a practice that can enhance communication among researchers, practitioners, and the agencies and people they serve—reinforcers are typically classified as edible, sensory, tangible, activity, or social.

Edible Reinforcers

Researchers and practitioners have used bites of preferred foods, snacks, candy, and sips of drinks as reinforcers. One interesting and important use of edibles as reinforcers is in the treatment of chronic food refusal in children. For example, Riordan, Iwata, Finney, Wohl, and Stanley (1984) used "highly preferred food items" as reinforcers to increase the food intake of four children at a hospital treatment facility. The treatment program consisted of dispensing the high-preference food items (e.g., cereal, yogurt, canned fruit, ice cream) contingent on the consumption of a target food item (e.g., vegetables, bread, eggs).

Kelley, Piazza, Fisher, and Oberdorff (2003) used edible reinforcers to increase cup drinking by Al, a 3-year-old boy who had been admitted to a day treatment program for food refusal and bottle dependency. The researchers measured the percentage of trials in which Al consumed 7.5 ml of three different liquids from the cup. During baseline, when Al was praised if he consumed the drink, his consumption averaged 0%, 44.6%, and 12.5% of trials for orange juice, water, and a chocolate drink, respectively. During the positive reinforcement component of the cup-drinking intervention, each time Al consumed the drink, the therapist praised him (as was done in baseline) and delivered a level spoon of peaches (a preferred food) to his mouth. Al consumed all three beverages on 100% of the trials during the positive reinforcement condition.

Fahmie, Iwata, and Jann (2015) found that 12 individuals diagnosed with intellectual disabilities or autism and ranging in age from 5 to 22 years preferred 15-sec access to a single edible reinforcer (e.g., M&M, raisin, pretzel) to 15-sec access to a leisure item (e.g., music box, plush bear, maraca) and that edible reinforcers maintained responding somewhat longer than leisure-item reinforcers. Regarding the implications of these results for effective practice, Fahmie and colleagues noted:

> These advantages provide evidence in favor of including edible items in preliminary preference assessments. One additional advantage of edible reinforcers is that they are delivered more easily and efficiently, which may have an overall beneficial effect on training in the long run. Nevertheless, researchers (e.g., Ferrari & Harris, 1981; Rincover & Newsom, 1985) have emphasized the relative benefits of leisure stimuli during training, stating that leisure stimuli are more natural, are less deleterious to health (see also Behavior Analysis Certification Board, 2014, Guideline 4.10), and promote more active engagement. Of course, behavior analysts also are responsible for providing the most effective treatment (Behavior Analyst Certification Board, 2014, Guideline 2.09). This study provides some additional evidence to help balance these goals. (p. 342, BACB citations updated to most recent version).

Sensory Reinforcers

Various forms of sensory stimulation such as vibration (e.g., massager), tactile stimulation (e.g., tickles, strokes with a feather boa), flashing or sparkling lights, and music have been used effectively as reinforcers (e.g., Bailey & Meyerson, 1969; Ferrari & Harris, 1981; Gast et al., 2000; Hume & Crossman, 1992; Rincover & Newsom, 1985; Vollmer & Iwata, 1991).

Tangible Reinforcers

Items such as stickers, trinkets, school materials, trading cards, and small toys often serve as tangible reinforcers. An object's intrinsic worth is irrelevant to its ultimate effectiveness as a positive reinforcer. Virtually any tangible item can serve as a reinforcer. Remember Engelmann's (1975) kindergarten students who worked for yellow slips of paper!

Activity Reinforcers

When the opportunity to engage in a certain behavior functions as reinforcement, that behavior is an activity reinforcer. Activity reinforcers may be everyday activities (e.g., playing a board game, leisure reading, listening to music), privileges (e.g., lunch with the teacher, shooting baskets, first in line), or special events (e.g., a trip to the zoo).

McEvoy and Brady (1988) evaluated the effects of contingent access to play materials on the completion of math worksheets by three students with autism and behavior disorders. During baseline, the teacher told the students to complete the problems as best that they could and to either complete other unfinished assignments or "find something else to do" if they finished the worksheets before a 6-minute timing elapsed. No other prompts or instructions were given for completing the worksheets. The teacher praised the completion of the worksheets.

On the first day of intervention for each student, he was taken to another room and shown a variety of toys and play materials. The teacher told the student he would have approximately 6 minutes to play with the materials if he met a daily criterion for completing math problems. Figure 11.7 shows the results. During baseline, the rate at which all three students correctly completed problems was either low (Dicky) or highly variable (Ken and Jimmy). When contingent access to the play activities was introduced, each student's completion rate increased and eventually exceeded criterion levels.

Premack (1959) hypothesized that activity reinforcers can be identified by looking at the relative distribution of behaviors in a free operant situation. Premack believed that behaviors themselves could be used as reinforcers and that the relative frequency of behavior was an important factor in determining how effective a given behavior might be as a reinforcer if the opportunity to engage in the behavior is contingent on another behavior. The **Premack principle** states that making the opportunity to engage in a behavior that occurs at a relatively high free operant (or baseline) rate contingent on the occurrence of low-frequency behavior will function as reinforcement for the low-occurrence behavior. For a student who typically spends much more time watching TV than doing homework, a contingency based on the Premack principle (informally known as

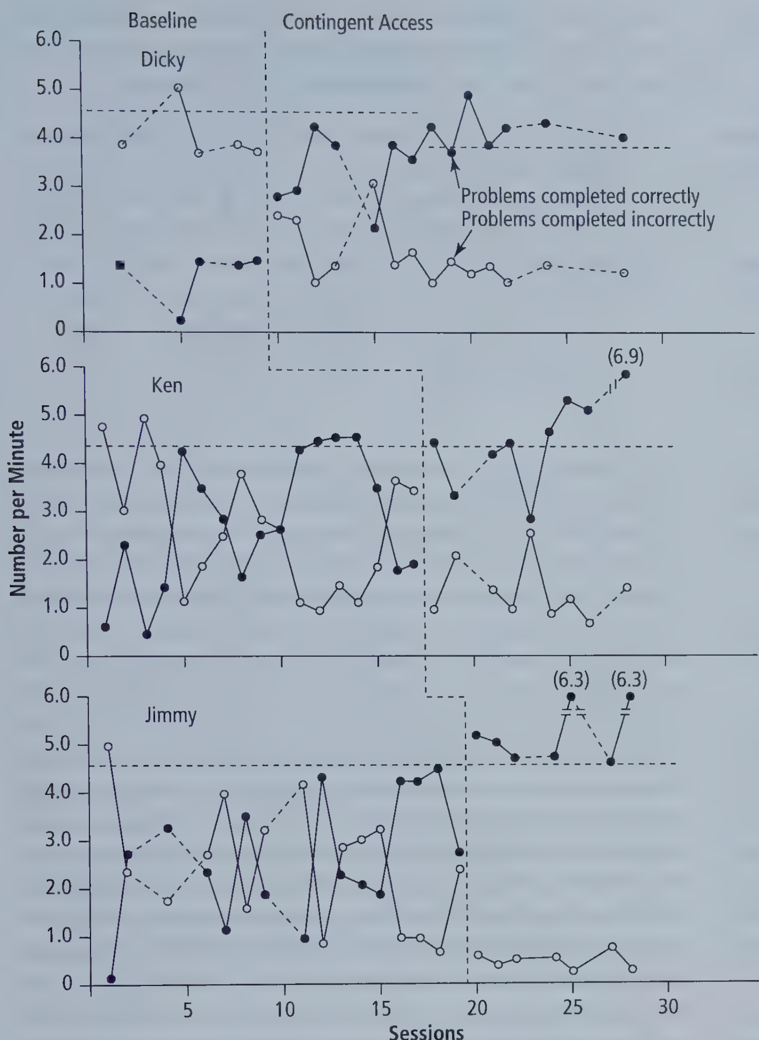


Figure 11.7 Number of math problems completed correctly and incorrectly per minute by three special education students during baseline and contingent access to play materials. Dashed horizontal lines indicate criteria.

From "Contingent Access to Play Materials as an Academic Motivator for Autistic and Behavior Disordered Children," by M. A. McEvoy and M. P. Brady, 1988, *Education and Treatment of Children, 11*, p. 15. Copyright 1998 by the Editorial Review Board of *Education and Treatment of Children*. Used by permission.

"Grandma's Law") might be, "When you have finished your homework, you can watch TV."

Building on Premack's concept, Timberlake and Allison (1974) proposed the **response-deprivation hypothesis** as a model for predicting whether access to one behavior (the contingent behavior) will function as reinforcement for another behavior (the instrumental response) based on the relative baseline rates at which each behavior occurs and whether access to the contingent behavior represents a restriction compared to the baseline level of engagement. Restricting access to a behavior presumably acts as a form of deprivation that serves as an EO, thus making the opportunity to engage in the restricted behavior an effective form of reinforcement (Allison, 1993; Iwata & Michael, 1994).

Iwata and Michael (1994) cited a series of three studies by Konarski and colleagues as demonstrating the veracity and applied implications of the response-deprivation hypothesis. In the first study, when students were given access to coloring (a high-probability behavior contingent on completing math problems (a low-probability behavior), they spent more time doing math, but only if the reinforcement schedule represented a restriction of the amount of time spent coloring compared to baseline (Konarski, Johnson, Crowell, & Whitman, 1980). The researchers found that a contingency in which students could earn *more time* coloring than they did in baseline for completing math problems was ineffective. These basic findings were reproduced in a subsequent study in which access to reading (or math, depending on the subject) was contingent on math (or reading) (Konarski, Crowell, Johnson, & Whitman, 1982). In the third study, Konarski, Crowell, and Duggan (1985) took the response-deprivation hypothesis a step further by examining the "reversibility of reinforcement" within subjects; that is, engaging in either of two activities—reading or math—could serve as reinforcement for increased performance in the other activity, in a response-deprivation condition for the contingent activity. Response deprivation for reading as the contingent response resulted in increases in math (instrumental response); conversely, response deprivation for math as the contingent response produced increases in reading. In all three studies, response restriction was the key factor in determining whether access to the contingent response would be reinforcing.

Iwata and Michael (1994) concluded that the collective results of Konarski and colleagues' studies illustrate each of three predictions based on the response-deprivation hypothesis

(assume the ratio of baseline rates of doing homework to watching TV is 1:2 in the following examples):

- Reinforcement of a low-rate target behavior when access to a high-rate contingent behavior is restricted below baseline levels (e.g., 30 minutes of homework gets access to 30 minutes of TV).
- Nonreinforcement of a low-rate behavior when access to a high-rate contingent behavior is not restricted below baseline levels (e.g., 30 minutes of homework gets access to 90 minutes of TV).
- Reinforcement of a high-rate target behavior when access to the low-rate behavior is restricted below baseline levels (e.g., 30 minutes of TV yields 5 minutes of homework).

Although recognizing that practitioners seldom design reinforcement programs to increase the rate of behaviors such as TV watching that already occur at high rates, Iwata and Michael (1994) noted that:

> There are a number of instances in which one may wish to produce highly accelerated performance (e.g., as in superlative academic or athletic performance that is good to begin with). In such cases, one need not find another activity that occurs at a higher rate to serve as reinforcement if one could arrange a suitable deprivation schedule with an activity that occurs at a relatively low rate. (p. 186)

As with all other descriptive categories of reinforcers, there is no a priori list that reveals what activities will or will not function as reinforcers. An activity that serves as effective reinforcement for one learner might have quite another effect on the behavior of another learner. For example, in Konarski, Crowell, and colleagues' (1982) study, access to math functioned as reinforcement for doing more reading for three students, whereas getting to read was the reinforcer for completing math problems for a fourth student. Many years ago, a classic cartoon brought home this crucial point very well. The cartoon showed two students dutifully cleaning the chalkboard and erasers after school. One student said to the other, "You're cleaning erasers for punishment!? I get to clean erasers as a reward for completing my homework."

Social Reinforcers

Physical contact (e.g., hugs, pats on the back), proximity (e.g., approaching, standing, or sitting near a person), attention, and praise are examples of events that often serve as social reinforcers. Adult attention is one of the most powerful and generally effective forms of reinforcement for children. The nearly universal effects of contingent social attention as reinforcement has led some behavior analysts to speculate that some aspects of social attention may entail unconditioned reinforcement (e.g., Gewirtz & Pelaez-Nogueras, 2000; Vollmer & Hackenberg, 2001).

The original experimental demonstrations and discovery of the power of adults' social attention as reinforcement for children's behavior took place in a series of four studies designed by Montrose Wolf and carried out by the preschool teachers at the Institute of Child Development at the University of Washington in the early 1960s (Allen, Hart, Buell, Harris, & Wolf, 1964; Harris, Johnston, Kelly, & Wolf, 1964; Hart, Allen, Buell, Harris, & Wolf, 1964; Johnston, Kelly, Harris, & Wolf, 1966). Describing those early studies, Risley (2005) wrote:

> We had never seen such power! The speed and magnitude of the effects on children's behavior in the real world of simple adjustments of something so ubiquitous as adult attention was astounding. Forty years later, social reinforcement (positive attention, praise, "catching them being good") has become the core of most American advice and training for parents and teachers—making this arguably the most influential discovery of modern psychology. (p. 280)

Because of the profound importance of this long-known but underused phenomenon, we describe a second study showing the effects of contingent attention as reinforcement for children's behavior. The first volume of the *Journal of Applied Behavior Analysis* included no fewer than seven studies building on and extending Wolf and colleagues' pioneering research on social reinforcement.[8] R. Vance Hall and colleagues conducted two of those studies. Like the Hall, Lund, and Jackson (1968) study, which introduced this chapter and from which we selected the example of a teacher's use of positive reinforcement with Robbie, the three experiments reported by Hall, Panyan, Rabon, and Broden (1968) continue to serve as powerful demonstrations of the effects of teacher attention as social reinforcement.

Participating in one of the experiments was a first-year teacher whose class of 30 sixth-graders exhibited such high rates of disruptive and off-task behaviors that the school principal described the class as "completely out of control." Throughout the study, Hall, Panyan, Rabon, and Broden (1968) measured teacher attention and student behavior during a continuous 30-minute observation period in the first hour of the school day. The researchers used a 10-second partial-interval observation and recording procedure to measure study behavior (e.g., writing the assignment, looking in the book, answering the teacher's question) and nonstudy behavior (e.g., talking out, being out of seat, looking out the window, fighting or poking a classmate). The observers also recorded the occurrence of teacher attention in each interval. Each instance of teacher verbal attention, defined as a comment directed to a student or group of students, was recorded with a "+" if it followed appropriate study behavior and with a "−" if it followed nonstudy behavior.

During baseline the class had a mean percentage of intervals of study behavior of 44%, and the teacher made an average of 1.4 comments following study behavior per session (see Figure 11.8). "Almost without exception those [comments] that followed study behavior were approving and those that followed nonstudy behavior were in the form of a verbal reprimand" (Hall, Panyan, Rabon, & Broden, 1968, p. 316). The level of study behavior by the class was 90% on one day when the helping teacher presented a demonstration lesson (see data points marked by solid arrows). On three occasions during baseline (data points marked by open arrows), the principal met with the teacher to discuss his organizational procedures in an effort to improve the students' behavior. These counseling sessions

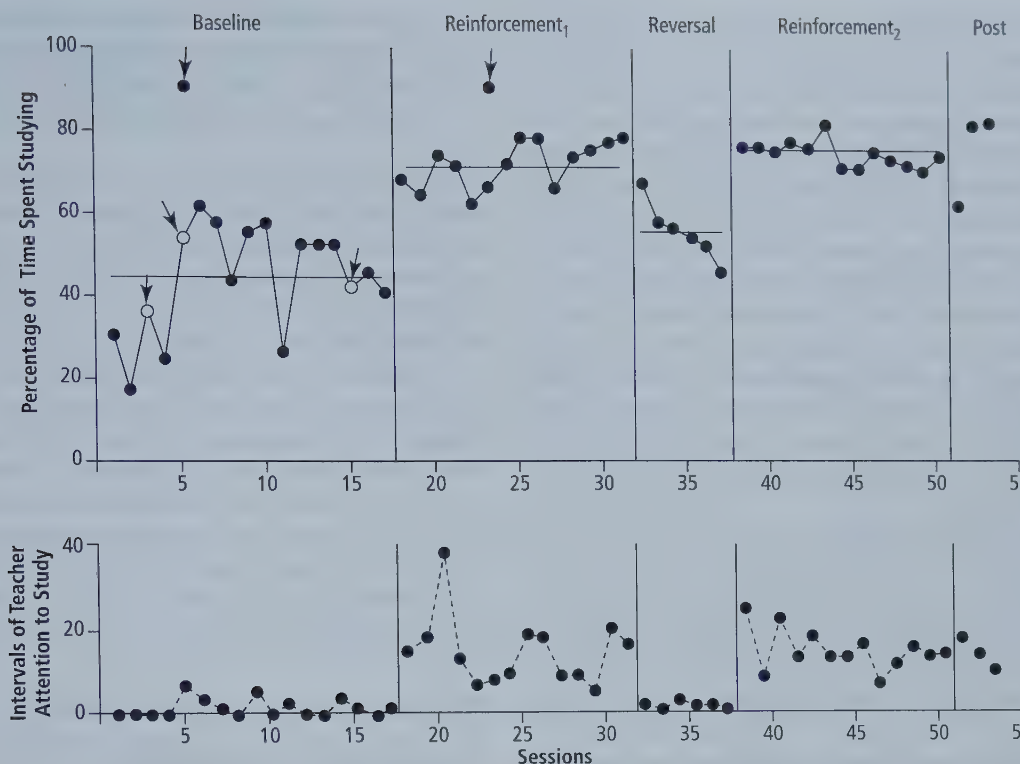


Figure 11.8 A record of class study behavior and teacher attention for study behavior during reading period in a sixth-grade classroom. Baseline = before experimental procedures; Reinforcement 1 = increased teacher attention for study; Reversal = removal of teacher attention for study; Reinforcement 2 = return to increased teacher attention for study. Post-follow up checks occurred up to 20 weeks after termination of the experimental procedures.

From "Instructing Beginning Teachers in Reinforcement Procedures Which Improve Classroom Control," by R. V. Hall, M. Panyan, D. Rabon, and M. Broden, 1968, *Journal of Applied Behavior Analysis*, *1*, p. 317. Copyright by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

resulted in the teacher writing all assignments on the board (after the first meeting) and changing the seating chart (after the third meeting). Those changes had no apparent effect on the students' behavior.

Prior to the first day of the reinforcement condition, the teacher was shown baseline data on the class study behavior and the frequency of teacher attention following study behavior. The teacher was instructed to increase the frequency of positive comments to students when they were engaged in study behavior. After each session during this condition, the teacher was shown data on the level of class study behavior and the frequency of his comments that followed study behavior. During the first reinforcement phase, teacher comments following study behavior increased to a mean frequency of 14.6, and the mean level of study behavior was 72%. The teacher, principal, and data collectors reported that the class was under better control and that noise had decreased significantly.

During a brief return of baseline conditions, the teacher provided "almost no reinforcement for study behavior," and a sharp downward trend in class study behavior was observed. The teacher, principal, and data collectors all reported that disruptive behavior and high noise levels had returned. The reinforcement conditions were then reinstated, which resulted in a mean frequency of 14 teacher comments following study behavior, and a mean level of 76% of intervals of study behavior.

IDENTIFYING POTENTIAL REINFORCERS

> *In the laboratory, we had learned to use a simple test: Place a candy in the palm of our hand, show it to the child, close our fist fairly tightly around the candy, and see if the child will try to pull away our fingers to get at the candy. If he or she will do that, even against increasingly tightly held fingers, the candy is obviously a reinforcer.*
>
> —Murray Sidman (2000, p. 18)

The ultimate success of many behavior change programs requires an effective reinforcer that the practitioner or researcher can deliver to increase future occurrences of the desired behavior targeted for improving. Fortunately, identifying effective and accessible reinforcers for most learners is relatively easy. Sidman (2000) described a quick and simple method for determining whether candy would likely function as a reinforcer. However, every stimulus, event, or activity that might function as a reinforcer cannot be held in the palm of the hand.

Identifying robust and reliable reinforcers for learners with severe and/or multiple disabilities may pose a major challenge. Although many common events (e.g., praise, music, free time, tokens) serve as effective reinforcers for most people, these stimuli do not function as reinforcers for all learners. Time, energy, and resources would be lost if planned interventions

were to fail because a practitioner used a presumed, instead of an actual, reinforcer.

Also, reinforcer preferences shift, and the transitory and idiosyncratic nature of preference has been reported repeatedly in the literature (Bishop & Kenzer, 2012; Carr, Nicholson, & Higbee, 2000; DeLeon et al., 2001; Ortiz & Carr, 2000; Shapiro, Kazemi, Pgosjana, Rios, & Mendoza, 2016). Preferences may change with the person's age, interest level, time of day, social interactions, deprivation and satiation levels, and the presence of establishing operations (EOs) (Gottschalk, Libby, & Graff, 2000; Whitehouse, Vollmer, & Colbert, 2014). What a teacher asks in September to determine preferences may have to be repeated a month later or a day later. Likewise, a therapist who asks a client what is reinforcing during a morning session may find that this stimulus is not preferred in an afternoon session (Bishop & Kenzer, 2012).

After reviewing 13 published studies that evaluated preferences and reinforcers for people with profound multiple disabilities, Logan and Gast (2001) concluded that preferred stimuli do not always function as reinforcers, and preferred stimuli at one point in time changed later. Additionally, individuals with behavioral or severe to profound developmental disabilities may engage in activities for such a limited time, or shift preferences within teaching sessions, that it is difficult to clearly determine whether a stimulus change is a reinforcer (Bishop & Kenzer, 2012).

To meet the challenge of identifying effective reinforcers, researchers and practitioners have developed a variety of procedures that fall under the twin headings of stimulus preference assessment and reinforcer assessment. Stimulus preference assessment and reinforcer assessment are often conducted in tandem to determine if a stimulus identified as a reinforcer actually functions as one (Kang et al., 2013; Lee, Yu, Martin & Martin, 2010; Piazza, Fisher, Hagopian, Bowman, & Toole, 1996; Whitehouse et al., 2014). Piazza et al. (1996) state:

> During preference assessments, a relatively large number of stimuli are evaluated to identify preferred stimuli. The reinforcing effects of a small subset of stimuli (i.e., the highly preferred stimuli) are then evaluated during reinforcer assessment. Although the preference assessment is an efficient procedure identifying potential reinforcers from a large number of stimuli, it does not evaluate the reinforcing effects of the stimuli. (pp. 1–2)

Stimulus preference assessment identifies stimuli that are likely highly preferred and thus likely to serve as reinforcers. As Kelley, Shillingsburg, and Bowen (2016) state, "The value of preference assessments is grounded in the extent to which the identified stimuli ultimately function as reinforcers for socially important behaviors" (p. 394).

Reinforcer assessment, in contrast, puts the highly preferred, potential reinforcers to a direct test by presenting them contingent on occurrences of a behavior and measuring performance effects (Fisher, Piazza, Bowman, & Amari, 1996). In this section we describe a variety of techniques developed by researchers and practitioners for conducting stimulus preference assessments and reinforcer assessment (see Figures 11.9 and 11.10). These methods form a continuum of approaches ranging from simple and quick to more complex and time-consuming.

Stimulus Preference Assessment

Stimulus preference assessment (SPA) refers to a variety of procedures used to determine (a) the stimuli that the person differentially selects, (b) the relative hierarchical preference value of those stimuli (high preference to low preference), (c) the conditions under which those preference values change when task demands, deprivation states, or schedules of reinforcement are modified, and (d) whether highly preferred items ultimately serve as effective reinforcers. Generally speaking, stimulus preference assessment is conducted using a three-step process: (1) A large pool of stimuli that might serve as reinforcers is gathered, (2) those stimuli are presented to the target person systematically to identify preference, and (3) high- and sometimes low-preference items are "tested" experimentally to determine the conditions under which they serve as reinforcers (Livingston & Graff, 2018). It is essential for practitioners to narrow the initial field of possible preferred stimuli to those that have good

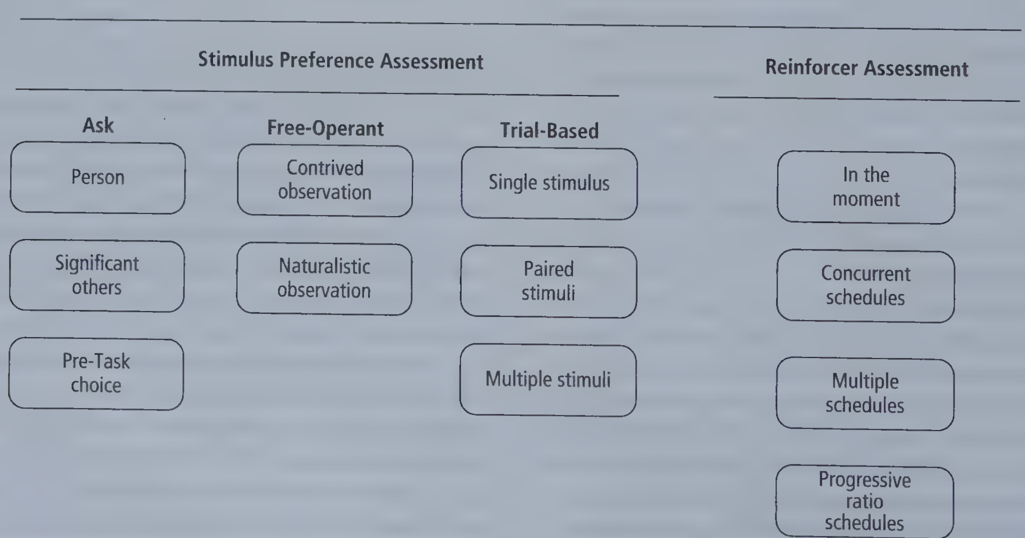


Figure 11.9 Stimulus preference assessment and reinforcer assessment methods for identifying potential reinforcers.

Figure 11.10 Commonly used stimulus preference assessments.

Name of Assessment (Citation)	Description of Assessment
Single-stimulus (SS) (Pace et al., 1985)	Across a series of trials, stimuli are presented one at a time. Approach responses (e.g., moving hand or body toward the item) are recorded. Preference hierarchies are established by calculating the percentage of approach responses per stimulus.
Paired-stimulus (PS) (Fisher et al., 1992)	Across a series of trials, stimuli are presented two at a time; individuals can approach (i.e., select) only one item on a trial. Approach responses are recorded. Preference hierarchies are established by calculating the percentage of approach responses per stimulus.
Multiple-stimulus-without-replacement (MSWO) (DeLeon & Iwata, 1996)	At the start of each session, multiple stimuli are placed in front of the individual, who can select one. Approach responses are recorded. The selected item is not replaced, and the positions of the remaining stimuli are changed. Then, the individual selects from the remaining items. Continue in this manner until all items have been selected or the individual stops selecting items. Typically, several sessions are conducted. Preference hierarchies are established by calculating the percentage of approach responses per stimulus across all sessions.
Brief Free Operant (FO) (Roane et al., 1998)	Multiple stimuli are placed on a tabletop, and participants are free to engage with any of the items for 5 min. Duration of engagement with each object (e.g., manipulating objects) is measured. Preference hierarchies are established by ranking items according to the duration of object manipulation for each stimulus.

From "Assessing Preferences of Individuals with Developmental Disabilities: A Survey of Current Practices," by R. B. Graff and A. M. Karsten, 2012a, *Behavior Analysis in Practice 5*(2), p. 38. Used with permission.

odds of functioning as reinforcers; otherwise, the assessment process will be protracted and intervention delayed.

Methodologically, SPA can be conducted using at least five forms: single-stimulus (SS), paired-stimulus (PS), free operant (FO), multiple-stimulus with item replacement (MSW), and multiple-stimulus without item replacement (MSWO). Each SPA method has advantages or disadvantages with respect to bias, assessment time required, likelihood of identifying reinforcers, and the chance that the assessment activity itself may evoke problem behavior (see Figure 11.11) (Karsten, Carr, & Lepper, 2011).

As illustrated in Figure 11.9 these five methodological variations can be clustered into three categories: (1) ask the person (or significant others) to identify preferred stimuli; (2) employ a free operant procedure by observing the person approaching, interacting, or engaging with various stimuli; and (3) measure responses to trial-based pairs or multiply presented stimuli by replacing or not replacing the stimulus after preference selection.

In choosing which method to use, practitioners must balance three competing perspectives: (a) gaining the maximum amount of preference assessment data within the least amount of time, but with false positives (i.e., believing a stimulus is preferred when it is not); (b) conducting a more extensive and time- and labor-intensive assessment that will delay intervention, but may yield more conclusive results; and (c) assessing the skill level of colleagues conducting the stimulus preference assessment and any initial or recurrent training needs they have

Figure 11.11 Assets and barriers associated with stimulus preference assessments.

Method	Assets	Potential barriers
MSWO	Likely to identify multiple reinforcers in minimal time	Positional bias, limited to smaller tabletop items and fewer items; requires more time than FO assessment
PS	Likely to identify multiple reinforcers, accommodates larger tabletop items and a greater number of items	Positional bias, requires more time than MSWO and FO assessments
SS	Likely to identify multiple reinforcers, accommodates larger items and activities	False positive results, less likely to identify relative preferences than MSWO and PS methods except when duration of engagement is also measured
FO	Less likely to evoke problem behavior, requires minimal time, accommodates larger items and activities	Less likely to identify multiple reinforcers than other methods except when assessment is repeated without most preferred item

Note: MSWO = multipe stimulus without replacement; PS = paired stimulus; SS = single stimulus; FO = free operant.

From "Description of a Practitioner Model for Identifying Preferred Stimuli with Individuals with Autism Spectrum Disorders," by A. M. Karsten, J. E. Carr, and T. L. Lepper, 2011, *Behavior Modification, 35*(4), p. 350. Copyright by Sage Publication. Reprinted by permission.

(Leaf et al., 2015). To assist analysts and practitioners with deciding which of the five methodological variations to employ, Karsten et al. (2011) proposed a three-step model that progresses from (a) choosing the SPA that is likely to yield an accurate hierarchy of high-preference stimuli, (b) modifying that SPA approach based on actual data collected during the SPA, and (c) verifying that the high-preference stimuli function as reinforcers.

Still, a practical roadblock may affect the implementation of the best of intended SPA models. Graff and Karsten (2012a) stated that upwards of one half of the respondents to their SPA questionnaire reported that they had little to no experience with conducting SPAs. Given that conducting some level of a SPA has become the gold standard of best practice for identifying highly preferred stimuli and ultimately reinforcers (Leaf et al., 2018), it seems important that training programs at the pre- and in-service levels address this shortcoming. Deliperi, Vladescu, Reeve, Reeve, and DeBar (2015) state the issue succinctly:

> Due to the reliance on reinforcement in interventions for consumers with disabilities, assessing the preferences of these individuals is an important part of effective programming. The failure to identify stimuli that function as reinforcers may compromise skill acquisition and behavior reduction programs. Therefore, identifying and implementing effective training procedures are essential to increasing the likelihood that staff will conduct preference assessments and use data from these assessments to guide their selection of consequences during subsequent programming. (p. 324)

Box 11.2 discusses a number of questions associated with training practitioners to conduct stimulus preference assessments.

BOX 11.2

Training Practitioners to Conduct Stimulus Preference Assessments (SPA)

Identifying reinforcers is key to a successful behavior change program. But busy, inexperienced, and inadequately trained preservice or in-service practitioners, who now also lack time and resources, are not likely to engage in conducting stimulus preference assessments without training from an experienced analyst (Graff & Karsten, 2012a). Although not intended to be exhaustive, the following frequently asked questions may provide a starting point to address stimulus preference assessment training issues for inexperienced, novice, or poorly trained preservice practitioners within either an individual or group-training context.

For SPAs to Be Conducted, What Skills Must Be Taught?

There appears to be consensus in the literature that practitioners need three main skills: (1) identifying stimulus items for the assessment; (2) conducting the SPA with a confederate or actual consumer; and (3) scoring and interpreting the results of the assessment to determine high- and low-preference stimuli. (Source: Deliperi et al., 2015)

Can Inexperienced Practitioners Learn the Skills Necessary to Conduct Reliable, Accurate, and Worthwhile Stimulus Preference Assessments?

Yes. Although the methods might vary and depend on local available resources, broadband Internet connectivity, and skill level of the trainer, it is quite possible—even desirable—to train (or retrain) inexperienced practitioners to use SPA. Doing so will give the teacher, parent, or caregiver direction for using those assessment data to further determine actual reinforcers that will affect future occurrences of target behaviors for the persons in their care. (Source: Ringdahl et al., 1997)

How Long Does It Take to Train an Inexperienced Person to Conduct a SPA?

The good news is that within several training sessions trainees can gain sufficient skill to conduct an SPA, especially if the training provides combinations of video models, scripts, voice-overs, practice, feedback, and supervision. (Sources: Bovi, Vladescu, DeBar, Carroll, & Sarokoff, 2017; Deliperi et al., 2015; Weldy, Rapp, & Capocosa, 2014)

Will Warranty Work Be Required? Does the Training Stick?

Whether future "booster" or "refresher" sessions will be required to maintain skills depends on a number of variables. For example, if novice practitioners only use SPA a few times per year, you can expect loss of initial skill acquisition. If, however, they use SPA regularly, it is more likely that initial training will be maintained, especially if they receive follow-up feedback and have access to review materials (prior training videos, embedded directions, etc.). In these cases, as the behavior analyst trainer, you would have more confidence that initial skills would be maintained. (Source: Deliperi et al., 2015)

Does SPA Training Only Work for "Ask" or "Free Operant" Categories of SPA?

No. A number of researchers have shown that analysts can train inexperienced practitioners to use *ask, free operant*, and *trial-based* variations of SPA. The complexity of the SPA methodology is not a determining or self-limiting factor in training. Inexperienced participants can learn a variety of SPA methods. (Sources: DeLeon & Iwata, 1996; Fisher et al., 1992; Hanley, Iwata, Roscoe, 2006; Pace, Ivancic, Edwards, Iwata, & Page, 1985)

Will Brief Written Instructions Alone—Like Those Found in a Method Section of a Journal Article—Be Sufficient for Training?

Probably not. Simply providing an opportunity for inexperienced practitioners to read the *method section* of a journal article on SPA is not likely to be successful. A number of researchers have demonstrated that written instructions alone are not robust enough to bring inexperienced practitioners up to a standard whereby they can conduct an SPA without supervision. Additional and supplementary materials are needed to "enhance" written instructions. These enhancements might include providing pictures, video models, role playing, voiceovers, feedback, and so on. (Sources: Graff & Karsten, 2012b; Lavie and Sturmey, 2002; Ramon, Yu, Martin, & Martin, 2015; Roscoe & Fisher, 2008; Shapiro et al., 2016).

If I Develop a Self-instruction Manual, Will That Work?

It depends. Although self-instruction manuals can be an improvement over brief, written methodological synopses excerpted from journals, the scope and sequence of the manual would be of primary importance. Otherwise, your trainee might be left with the equivalent of "written instructions alone," which has not been shown to be effective. Using a self-instruction manual alone in this manner would be tantamount to a "train and hope" process. However, if the manual provides introductory content, step-by-step directions, a procedural checklist, exercises, an answer key, and methods for recording, summarizing, and interpreting SPA data, the manual's utility is enhanced considerably. If the trainer uses feedback from the trainees as a way to improve the accuracy and completeness of the manual, the manual would evolve over time into a more effective document. (Source: Ramon et al., 2015)

Can I Use Video, Multimedia, or a Smartphone App to Teach SPA?

Yes. Video methods have been shown to be effective, especially when combined as a "package approach," including feedback, embedded text, rehearsal, and supervision. Also, trainers can use smartphone apps (e.g., Plickers.com) to conduct basic SPA formative assessments with groups of students. (Sources: Hansard & Kazemi, 2018; Lavie & Sturmey, 2002; Radley, Dart, Battaglia, & Ford, 2018)

I Have a Number of Practitioners to Teach. Can I Teach Within a Group Context?

Yes. Teaching practitioners within a small-group context can increase training efficiency, without loss of effect. (Source: Weldy et al., 2014)

I Live in a Rural/Remote Area. Can I Use Distance Learning with My Staff?

Yes. One of the remarkable benefits of high-speed, broadband Internet connections is that telecommunication links can be established and operated remotely. Training staff in distant locations appears to be at least as comparable to in-person training with respect to initial skill acquisition and maintenance, even without in-person contact or on-site trainers. (Source: Higgins, Luczynski, Carroll, Fisher, & Mudford, 2017)

Choosing a method to teach inexperienced practitioners rests with the trainer's assessment of current staff resources (e.g., time, physical facilities, exigent circumstances, location) and consumer needs. Although a number of methods can be effective, depending on individual or group instruction, trainers would be advised to follow a "model-lead-test" sequence to ensure initial training to mastery. If necessary, be prepared to provide follow-up "booster" sessions to ensure maintenance and retention. Do not assume that "once taught is forever learned," especially if the practitioner has few opportunities to implement SPA after initial training.

Asking About Stimulus Preferences

A person's preference for various stimuli might be determined by merely asking what she likes. Asking can greatly reduce the time needed for more intrusive stimulus preference assessment, and it often yields information that can be integrated within an intervention program. Several variations of asking exist: asking the target person, asking significant others in the person's life, or offering a pretask choice assessment.

Asking the Target Person. A straightforward method for determining stimulus preference is to ask the target person what she likes. Typical variations include asking open-ended questions, providing the person with a list of choices or asking her to rank-order a list of choices.

- ***Open-ended questions.*** Depending on the learner's language abilities, an open-ended assessment of stimulus preference can be done orally or in writing. The person may be asked to name preferences among general categories of reinforcers—for example, "What do you like to do in your free time?" "What are your favorite foods and drinks?" "Are there any types of music or performers whose music you like?" An open-ended assessment can be accomplished simply by asking the learner to list as many favorite activities or items as possible. She should list not only everyday favorite things and activities but also special items and activities. Figure 28.15 is simply a sheet with numbered lines on which family members identify potential rewards they would like to earn by completing tasks on contingency contracts.

- ***Choice format.*** This format could include asking questions such as the following: "Which would you do a lot of hard work to get?" "Would you rather get things to eat, like chips, cookies, or popcorn, or get to do things like art projects, play computer games, or go to the library?" (Northup, George, Jones, Broussard, & Vollmer, 1996, p. 204)
- ***Rank-ordering.*** The learner can be given a list of items or stimuli and instructed to rank-order them from most to least preferred.

For learners with limited language skills, pictures of items, icons, or, preferably, the actual stimuli can be presented (Clevenger & Graff, 2005). For example, a teacher, while pointing to an icon, might ask a student, "Do you like to drink juice, use the computer, ride the bus, or watch TV?" Students simply nod yes or no.

Surveys have been developed to assess students' preferences. For example, elementary school teachers might use the *Child Reinforcement Survey*, which includes 36 rewards in four categories: edible items (e.g., fruit, popcorn), tangible items (e.g., stickers), activities (e.g., art projects, computer games), and social attention (e.g., a teacher or friend saying, "I like that") (Fantuzzo, Rohrbeck, Hightower, & Work, 1991). Other surveys are the *School Reinforcement Survey Schedule* for students in grades 4 through 12 (Holmes, Cautela, Simpson, Motes, & Gold, 1998) and the *Reinforcement Assessment for Individuals with Severe Disabilities* (Fisher et al., 1996).

Although asking for personal preferences is relatively uncomplicated, the procedure is not entirely foolproof with respect to confirming that a preferred choice will later serve as a reinforcer (Whitehouse et al., 2014). "Poor correspondence between verbal self-reports and subsequent behavior has been long noted and often demonstrated" (Northup, 2000, p. 335). Although a child might identify watching cartoons as a preferred event, watching cartoons may function as a reinforcer only when the child is at home on Saturday mornings, but not at Grandma's house on Sunday night.

Further, surveys may not differentiate accurately between what children claim to be high-preference and low-preference items for reinforcers. Northup (2000) found that preferences of children with attention-deficit hyperactivity disorder (ADHD) did not rise beyond chance levels when survey results were later compared to reinforcer functions. "The relatively high number of false positives and low number of false negatives again suggest that surveys may more accurately identify stimuli that are not reinforcers than those that are" (p. 337). Merely asking children their preferences once might lead to false positives (i.e., children may choose an event or stimulus as a reinforcer, but it may not be reinforcing).

Asking Significant Others. A pool of potential reinforcers can be obtained by asking parents, siblings, friends, or caregivers to identify the activities, items, foods, hobbies, or toys that they believe the learner prefers. For example, the *Reinforcer Assessment for Individuals with Severe Disabilities (RAISD)* is an interview protocol that asks caregivers to identify preferred stimuli across visual, auditory, olfactory, edible, tactile, and social domains (Fisher et al., 1996). Significant others then rank-order the selected preferences based on likely high-versus low-preference items. Finally, significant others are asked to identify the conditions under which they predict that specific items might function as reinforcers (e.g., cookies with milk versus just cookies alone). As a cautionary note, however, analysts should be aware that stimuli identified as highly preferred by a significant other's opinion *alone* may not always function as an effective reinforcer. Fisher et al. (1996) remind us: "Caregiver opinion alone is insufficient for accurate reinforcer identification" (p.14). However, when caregiver opinion is combined with a structured interview that provides prompts, cues, and a broader array of potential reinforcers for the significant other to consider, the caregiver report can be a helpful supplement to other methods.

Offering a Pretask Choice. In this method, the practitioner asks the participant to choose what he wants to earn for doing a task. The participant then chooses one item from two or three options presented (Piazza et al., 1996). All of the stimuli presented as pretask choices will have been identified as preferred stimuli by other assessment procedures. For instance, a teacher might make the following statement: "Robyn, when you finish your math problems, you may have 10 minutes to play *Battleship* with Martin, read quietly, or help Ms. Obutu prepare the social studies poster. Which activity do you want to work for?" A learner's choice of a consequence will not necessarily be a more effective reinforcer than one selected by the researcher or practitioner (Smith, Iwata, & Shore, 1995).

Free Operant Observation

The activities that a person engages in most often when able to choose freely from among behaviors will often serve as effective reinforcers when made contingent on engaging in low-probability behaviors. Observing and recording what activities the target person engages in when she can choose during a period of unrestricted access to numerous activities is called *free operant observation.* A total duration measure of the time the person engages with each stimulus item or activity is recorded. The longer the person engages with an item, the stronger the inference that the item is preferred.

Procedurally, the person has unconstrained and simultaneous access to a predetermined set of items or activities or to the materials and activities that are naturally available in the environment. There are no response requirements, and all stimulus items are available and within the person's sight and reach. An item is never removed after engagement or selection. According to Ortiz and Carr (2000) and Karsten et al. (2011), free operant responding is less likely to produce problem behavior during the assessment that might otherwise be observed if a stimulus is removed. Free operant observations can be contrived or conducted in naturalistic settings.

Contrived Free Operant Observation. Practitioners use contrived observation to determine whether, when, how, and the extent to which the person engages with each of a predetermined set of activities and materials. The observation is contrived because the researcher or practitioner "salts" the environment with a variety of items that may be of interest to the learner.

Figure 11.12 Number of minutes Mike spent engaged in activities during 2 hours of free time after school.

Activity	Mon	Tue	Wed	Thus	Fri	Total
Leisure reading	—	10	—	10	—	20
Watch TV	35	50	60	30	30	205
Phone with friends	15	15	10	20	10	70
Play video games	70	45	40	60	80	295
Play with constuction toy	—	—	10	—	—	10
Minutes observed	**120**	**120**	**120**	**120**	**120**	**600**

Free operant assessment presupposes that the person has had sufficient time to move about and explore the environment and has had the chance to experience each of the stimuli, materials, or activities. Just prior to the free operant observation period, the learner is provided brief noncontingent exposure to each item. All of the items are then placed within view and easy access to the learner, who then has the opportunity to sample and choose among them freely. Observers record the total duration of time that the learner engages with each stimulus item or activity.

Naturalistic Free Operant Observation. Naturalistic observations of free operant responding are conducted in the learner's everyday environment (e.g., playground, classroom, home). As unobtrusively as possible, the observer notes how the learner allocates his time and records the number of minutes the learner devotes to each activity. For instance, Figure 11.12 shows how a teenager, Mike, distributed his time during 2 hours of free time each day after school. Mike's parents collected these data by keeping a chart of the total number of minutes their son was engaged in each activity. The summary chart for the week shows that Mike played computer video games, watched television, and talked on the phone to his friends every day. On two different days Mike spent 10 minutes reading a library book, and he played with a new construction toy for a brief time on Wednesday. Two activities—watching television and playing video games—occurred the most often and for the longest duration. If Mike's parents wanted to apply the Premack principle introduced earlier in this chapter to increase the amount of time he spends reading for pleasure or playing with the construction toy (i.e., low-probability behaviors), they might make watching television or playing video games (i.e., high-probability behaviors) contingent on a certain amount of time spent leisure reading or playing with the construction toy.

Trial-based Methods

In trial-based methods of stimulus preference assessment, stimuli are presented to the learner in a series of tests, and the learner's responses to the stimuli are measured as an index or hierarchy of preference. One or more of three measures of the learner's behavior are recorded in trial-based stimulus preference assessment: approach, contact (DeLeon & Iwata, 1996), and engagement with the stimulus (DeLeon, Iwata, Conners, & Wallace, 1999; Hagopian, Rush, Lewin, & Long, 2001; Roane, et al., 1998). *Approach* responses typically include any detectable movement by the person toward the stimulus (e.g., eye gaze, head turn, body lean, hand reach), a *contact* is tallied each time the person touches or holds the stimulus, and *engagement* is a measure of the total time or percentage of observed intervals in which the person interacts with the stimulus (e.g., in which the person held a massager against her leg). An assumption is made that the more frequently the person approaches, touches or holds, or engages with a stimulus, the more likely the stimulus is preferred. As DeLeon and colleagues (1999) stated, "duration of item contact is a valid index of reinforcer value" (p. 114).

Preferred stimuli are typically labeled along a hierarchy from high-preference (HP) to low-preference (LP) stimuli based on predetermined criteria (e.g., stimuli chosen 80% or more of the time are HP) (Carr et al., 2000; Northup, 2000; Pace et al., 1985; Piazza et al., 1996). Implicit, but testable, assumptions are that a highly preferred stimulus will serve as a reinforcer and that preference stability is necessary before treatment can begin. Although these assumptions do not always hold (Higbee, Carr, & Harrison, 2000; Verriden & Roscoe, 2016), they have proven to be efficient and workable assumptions with which to begin.

The many variations of trial-based stimulus preference assessment can be grouped by presentation method as single stimulus (successive choice), paired stimuli (forced choice), and multiple stimuli with or without replacement.[9]

Single Stimulus. A single-stimulus presentation method, also called a *successive choice* or *single-stimulus engagement preference* method, represents the most basic assessment available for determining preference. Simply stated, a stimulus is presented by a trained clinician or teacher, and the person's reaction to, or engagement with, the stimulus is noted. Engagement behaviors include touching, playing with, looking at, or holding the stimulus for a predetermined time. Presenting one stimulus at a time "may be well suited for individuals who have difficulty selecting among two or more stimuli" (Hagopian et al., 2001, p. 477).

To conduct a single-stimulus assessment, target stimuli across all sensory systems (i.e., visual, auditory, vestibular, tactile, olfactory, gustatory, and multisensory) are made available to the participant for a brief period. The participant is allowed to sample the stimuli one at a time, engage with it briefly, and, if a toy, see how it works, before the formal assessment protocol begins. Next, the stimuli are presented one at a time in random order, and the person's reaction to and/or engagement with

each stimulus is recorded (Logan et al., 2001; Pace et al., 1985). Approach, avoidance, or rejection responses are recorded in terms of occurrence (yes or no), count (e.g., number of touches per minute), or duration (i.e., time spent engaged with an item). Each item in the array of proposed stimuli is presented sequentially. For example, a mirror might be presented to determine the duration of time the participant gazes into it, touches it, or rejects the mirror altogether (i.e., pushes it away). Each item is presented several times, and the order of presentation should be varied. Once all stimuli have been presented, a graphical hierarchy is constructed with the stimuli listed on the *x*-axis and the percentage of engagement on the *y*-axis. By analyzing the data with a bar graph, stimuli that are ranked the highest are presumed to be the preferred reinforcers.[10]

Boutot and DiGangi (2018) used a single-stimulus assessment procedure to identify preferred and nonpreferred toys by Dominic, a 5-month-old infant with Down syndrome. If Dominic looked at the adult-activated toy for longer than 5 seconds, changed facial expressions while looking at the object (e.g., smiled), or produced a marked bodily movement (kicked his legs or waved his arms), the toy was deemed preferred. Using this method, the researchers ultimately identified three toys and later employed them as reinforcers to increase the percentage of time Dominic sustained his head in an elevated position during tummy time over a "No Toy" condition (See Figure 11.13). Boutet and DiGangi (2018) concluded that the single-stimulus assessment procedure could be used in an efficient, probe-like manner to identify preferred stimuli and that those stimuli could be used by researchers and parents to improve important muscular skills for an infant with Down syndrome.

Paired Stimuli. Each trial in the paired-stimuli presentation method, also sometimes called the "forced choice" method, consists of the simultaneous presentation of two stimuli. The observer records which of the two stimuli the learner chooses. To conduct a paired-stimuli assessment, each stimulus is matched randomly with all other stimuli in the proposed group of stimuli to be compared (Fisher et al., 1992). The number in the group is at the discretion of the analyst. For example, Piazza and colleagues (1996) used 66 to 120 paired-stimuli trials to determine reinforcer preference. Data from a paired-stimuli assessment show how many times each stimulus is chosen. The stimuli are then rank-ordered from high to low preference. Items that received engagement, play, manipulation, or selection over 80% of the pairings are deemed high-preference items. Low-preference items were selected at the 50% level. Pace and colleagues (1985) found that paired-stimuli presentations yielded more accurate distinctions between high- and low-preference items than did single-stimulus presentations. Paired-stimulus preference assessment has the advantage of arranging all stimuli sequentially, thereby providing choices that otherwise would not be possible during a single-stimulus presentation format. Thus,

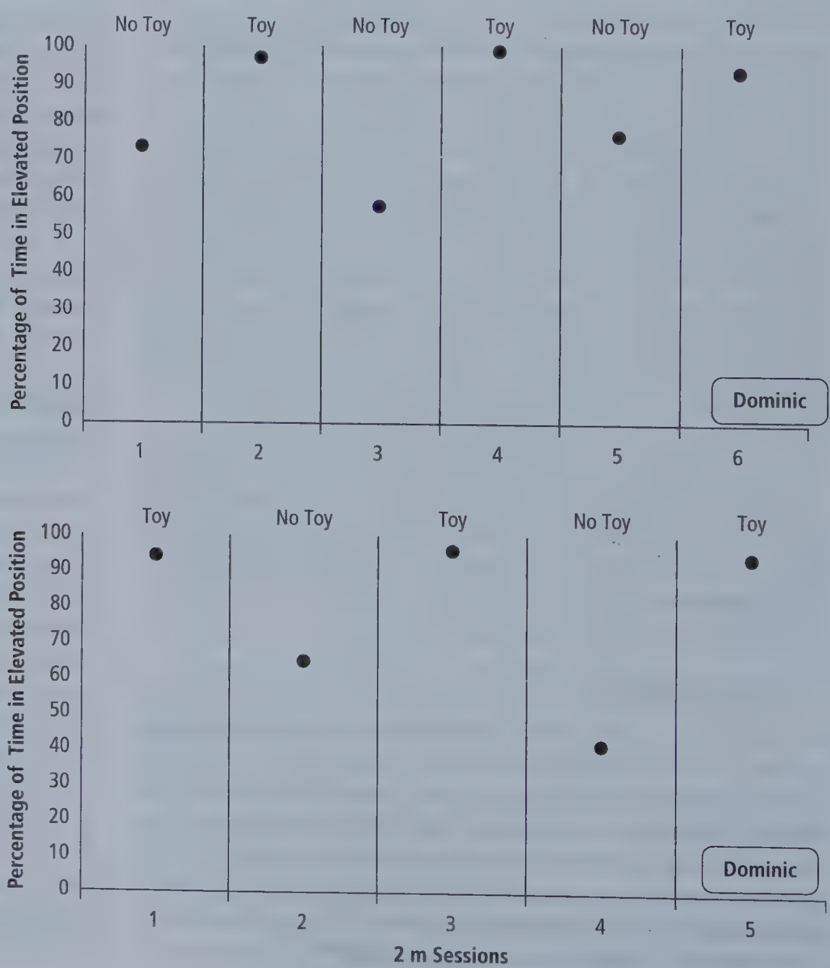


Figure 11.13 Percentage of each 2-minute session in which Dominic elevated his head above a 45° angle when the researcher (top panel) or parent (bottom panel) activated a preferred toy during tummy time or when no toy was present.

From "Effects of Activation of Preferred Stimulus on Tummy Time Behavior of an Infant with Down Syndrome and Associated Hypotonia," by E. A. Boutot and S. A. DiGangi, 2018, *Behavior Analysis in Practice*, *11*, p. 146. Copyright by the Association for Behavior Analysis International. Reprinted by permission.

the likelihood of ultimately identifying reinforcers is higher (Paclawskyj & Vollmer, 1995).

Because every possible pair of stimuli must be presented, paired-stimuli assessing is likely to take more time than the simultaneous presentation of a series of multiple stimuli. However, DeLeon and Iwata (1996) argued that ultimately the paired-stimuli method may be more time efficient because "the more consistent results produced by the PS method may indicate that stable preferences can be determined in fewer, or even single, sessions" (p. 520)[11].

Multiple Stimuli. The multiple-stimuli presentation method is an extension of the paired-stimuli procedure developed by Fisher and colleagues (1992). The participant to be assessed chooses a preferred stimulus from an array of three or more stimuli (Windsor, Piche, & Locke, 1994). By presenting multiple stimuli together, assessment time is reduced. For example, instead of presenting a series of trials consisting of all possible pairs of stimuli from a group of six stimuli and continuing until all pairs have been presented, all six stimuli are presented simultaneously.

The two major variations of the multiple-stimuli preference assessment are multiple stimuli with replacement (MSWI) and multiple stimuli without replacement (MSWO). The difference between the two lies in which stimuli are removed or replaced in preparation for the next trial after the participant indicates a preference among the displayed items. In the multiple stimuli with replacement procedure, the item chosen by the learner remains in the array and items that were not selected are replaced with new items. In the multiple stimuli without replacement procedure, the chosen item is removed from the array, the order or placement of the remaining items is structurally rearranged, and the next trial begins with a reduced number of items in the array (Brodhead, Al-Dubayan, Mates, Abel, & Brouwers, 2016).[12]

In any case, each trial begins by asking the person, "Which one do you want the most?" (Higbee et al., 2000), or instructing the person to "Choose one" (Ciccone, Graff, & Ahearn, 2005) and then continues until all items from the original array, or the gradually reducing array, have been selected. The entire sequence is usually repeated several times, although a single round of trials may identify stimuli that function as reinforcers (Carr et al., 2000).

The stimuli presented in each trial might be tangible objects themselves, pictures of the items, or verbal descriptions. Comparisons between conducting stimulus preference assessments with tangible objects themselves versus pictures of those objects have been mixed. Whereas Brodhead et al. (2016) have shown that video MSWO may be a suitable complement to tangible MSWI, Higbee, Carr, and Harrison (1999) found that the tangible objects produced greater variation and distribution of preferences than the picture objects did. Cohen-Almeida, Graff, and Ahearn (2000) found that the tangible object assessment was about as effective as a verbal preference assessment, but the clients completed the verbal preference assessment in less time.

In sum, MSWO assessment may identify stimuli that function as reinforcers, may require fewer sessions or less time to complete than other variations of preference assessments, yields a hierarchy of student preference, and appears to be suitable for a wide range of students due to fast administration and accuracy in identifying potential reinforcers (Brodhead et al., 2016).

DeLeon and Iwata (1996) used an adaptation of the multiple-stimuli and paired-stimuli presentations they described as a *brief stimulus assessment* to reduce the time needed to determine stimulus preference. Basically, in the brief stimulus assessment, once a particular stimulus item is chosen, that item is *not* returned to the array. Subsequent trials present a reduced number of items from which to choose (Carr et al., 2000; DeLeon et al., 2001; Roane et al., 1998). DeLeon and Iwata (1996) found that multiple stimuli without replacement identified preferred items in approximately half the time that a paired-stimulus comparison procedure did. According to Higbee and colleagues (2000), "With a brief stimulus preference procedure, practitioners have a method for reinforcer identification that is both efficient and accurate" (pp. 72–73).

Guidelines for Selecting and Using Stimulus Preference Assessments

Practitioners can combine assessment procedures to compare single versus paired, paired versus multiple, or free operant versus trial-based methods (Ortiz & Carr, 2000).[13] In day-to-day practice, brief stimulus presentations using comparative approaches might facilitate reinforcer identification, thereby speeding up interventions using those reinforcers. For teaching purposes and to launch instruction as soon as possible, Karsten et al. (2011) recommend using FO and/or MSWO options, as they are efficient and likely to provide sufficient information on a reinforcer hierarchy. Tung, Donaldson, and Kang (2017) also state that the FO procedure may be the method of choice when problem behavior is anticipated when tangible stimuli are removed during the assessment process.

In summary, an important goal of stimulus preference assessments is to identify stimuli that are most highly preferred and thus likely to function as reinforcers (Kelly, Roscoe, Hanley, & Schlichenmeyer, 2014). Each method for assessing preference has advantages and limitations with respect to identifying preferences (Karsten et al., 2011; Roane et al., 1998). Practitioners may find the following guidelines helpful when conducting stimulus preference assessments (DeLeon & Iwata, 1996; Gottschalk et al., 2000; Higbee et al., 2000; Ortiz & Carr, 2000; Roane et al., 1998; Roscoe, Iwata, & Kahng, 1999):

- Monitor the learner's activities during the time before the stimulus preference assessment session to be aware of EOs that may affect the results.
- Use stimulus preference assessment options that balance the cost–benefit of brief assessments (but the possible occurrence of false positives) with more prolonged assessments, which, while they may ultimately identify more highly preferred stimuli (and reinforcers), delay intervention.
- Balance using a stimulus preference method that may yield a ranking of preferred stimuli against an assessment method that occurs without rankings, but occurs more frequently, to counteract shifts in preference.
- Conduct a brief stimulus preference assessment with fewer items in an array (i.e., 3 instead of 5 or more) when assessment time is limited.
- Combine data from multiple assessment methods and sources of stimulus preference (e.g., asking the learner

and significant others, free operant observation, and trial-based methods).

- Verify that stimuli identified by any efficient SPA method ultimately serve that function as a reinforcer.
- Recognize that whereas stimulus preference may be stable for individuals over time (Kelley et al., 2016), the possibility exists that stimulus preference will shift, and further SPAs may be necessary.
- To facilitate SPA assessment with respect to time, consider using categories of stimuli. Edible categories might include chocolate, gummy treats, crunchy snacks, and so forth. Toy categories might include, for example, trucks, cars, and fire engines. It's likely that if SPA shows a preference for one item within a category, other unassessed items within that category might also serve as reinforcers, even if they are not initially formally assessed.

Reinforcer Assessment

The only way to tell whether or not a given event is reinforcing to a given organism under given conditions is to make a direct test.

—B. F. Skinner (1953, pp. 72–73)

Highly preferred stimuli may not always function as reinforcers (Higbee et al., 2000); even the candy that a child pried out of Sidman's hand may not have functioned as reinforcer under certain conditions. Conversely, least preferred stimuli might serve as reinforcers under some conditions (Gottschalk et al., 2000). The only way to know for sure whether a given stimulus serves as a reinforcer is to present it immediately following the occurrence of a behavior and note its effects on responding.

Reinforcer assessment refers to a variety of direct, data-based methods used to present one or more stimuli contingent on a target response and then measuring the future effects on the rate of responding. Researchers and practitioners have developed reinforcer assessment methods to determine the relative effects of a given stimulus as reinforcement under different and changing conditions and to assess the comparative effectiveness of multiple stimuli as reinforcers for a given behavior under specific conditions. Reinforcer assessment or analysis is often accomplished by presenting stimuli suspected of being reinforcers either "in the moment" or contingent on responding within concurrent, multiple, or progressive-ratio reinforcement schedules.[14]

In-the-Moment Reinforcer Analysis

In-the-moment reinforcer analysis is conducted by practitioners when a learner's response is followed immediately by the presentation of a stimulus change—in this case, a stimulus presumed to be a reinforcer—and the effect is noted on increases in the future occurrence of similar responses (Leaf et al., 2018). In practical applications, the analyst, with immediate access to numerous presumed reinforcers, makes an instantaneous judgment about which of those "reinforcers" is likely to function as such, and delivers that reinforcer to the learner. When deciding which reinforcer to deliver, the analyst considers the learner's current affect (smiling, frowning), previous interactions with that stimulus item (approaching, engaging, contacting), past data showing the effectiveness of that stimulus with respect to performance (improved or decreased), interactions with similar items in the past (furry dog, fleece-covered cat), and other past observational preferences (Leaf et al., 2012).

Leaf et al. (2018) compared formal paired-stimulus preference assessments with "in-the-moment reinforcer analysis" used to determine rates of learning a simple picture-labeling task by two preschool boys with autism. Procedurally, the teachers conducted a paired-stimulus preference assessment to determine the top three preferred items. During teaching, if the boys answered a question correctly, they received access to one of the top three items for 10 seconds. If they answered incorrectly, they received feedback that their response was incorrect. During the in-the-moment condition, the identical pool of paired-stimulus condition items was available, but the teachers had complete discretion about which item to deliver contingent on a correct response. They could deliver the same reinforcer during all learning trials or change the reinforcer after each trial. The determining factor was their moment-to-moment judgment based on the factors cited above.

The results of Leaf et al. showed no clear difference in skill acquisition for the boys—they learned equally well in both conditions, but the in-the-moment procedure was far more efficient with respect to the division of time spent in assessment versus teaching conditions. Leaf and colleagues concluded by stating: "When teaching individuals with learning difficulties, any incremental gain in speed of acquisition through formal preference assessments must be weighed against the reduced teaching time that remains after completing the assessments, which is a point that seems to be too often overlooked" (p. 23).

Concurrent Schedule Reinforcer Assessment

When two or more contingencies of reinforcement operate independently and simultaneously for two or more behaviors, a *concurrent schedule of reinforcement* is in effect. When used as a vehicle for reinforcer assessment, a concurrent schedule arrangement essentially pits two stimuli against each other to see which will produce the larger increase in responding when presented as a consequence for responding. If a learner allocates a greater proportion of responses to one component of the concurrent schedule over the other, the stimulus used as a contingent consequence for that component is the more effective reinforcer. Using a concurrent schedule in this way shows the relative effectiveness of high-preference (HP) and low-preference (LP) stimuli as reinforcers (Koehler, Iwata, Roscoe, Rolider, & O'Steen, 2005; Piazza et al., 1996).

Concurrent schedules may also be used to determine differences between *relative* and *absolute* reinforcement effects of stimuli. That is, will an LP stimulus now presented contingently in the absence of the HP stimulus serve as a reinforcer? Roscoe and colleagues (1999) used concurrent schedules to compare the effects of HP and LP stimuli as reinforcers for eight adults with developmental disabilities. Following the preference assessments, a concurrent schedule of reinforcement was established using the high-preference and low-preference stimuli. The target response was pressing either of two micro switch panels. Each panel was a different color. Pressing a panel would illuminate

a small light in the center of the panel. A training condition took place prior to baseline to establish panel pressing in the subjects' repertoires and to expose them to the consequences of responding. During baseline, pressing either panel resulted in no programmed consequences. During the reinforcement phase, an HP stimulus was placed on a plate behind one of the panels, and an LP stimulus was placed on a plate behind another panel. All responses to either panel resulted in the participant immediately receiving the item on the plate behind the respective panel (i.e., an FR 1 schedule of reinforcement).

Under the concurrent schedule of reinforcement that enabled participants to choose reinforcers on the same FR 1 schedule, the majority of participants allocated most of their responding to the panel that produced the HP stimulus as reinforcement (e.g., see results for Sean, Peter, Matt, and Mike on Figure 11.14). However, these same participants, when later presented with the opportunity to obtain LP stimuli as reinforcers on a single-schedule contingency (i.e., only one panel to push), showed increased levels of responding over baseline, similar to those obtained with the HP stimuli in the concurrent schedule. The study by Roscoe and colleagues (1999) demonstrated how concurrent schedules may be used to identify the relative effects of stimuli as reinforcers. The study also showed that the potential effects of a stimulus as a reinforcer may be masked or overshadowed when that stimulus is pitted against another stimulus on a concurrent schedule. In such cases, a potentially reinforcing stimulus might be abandoned prematurely.

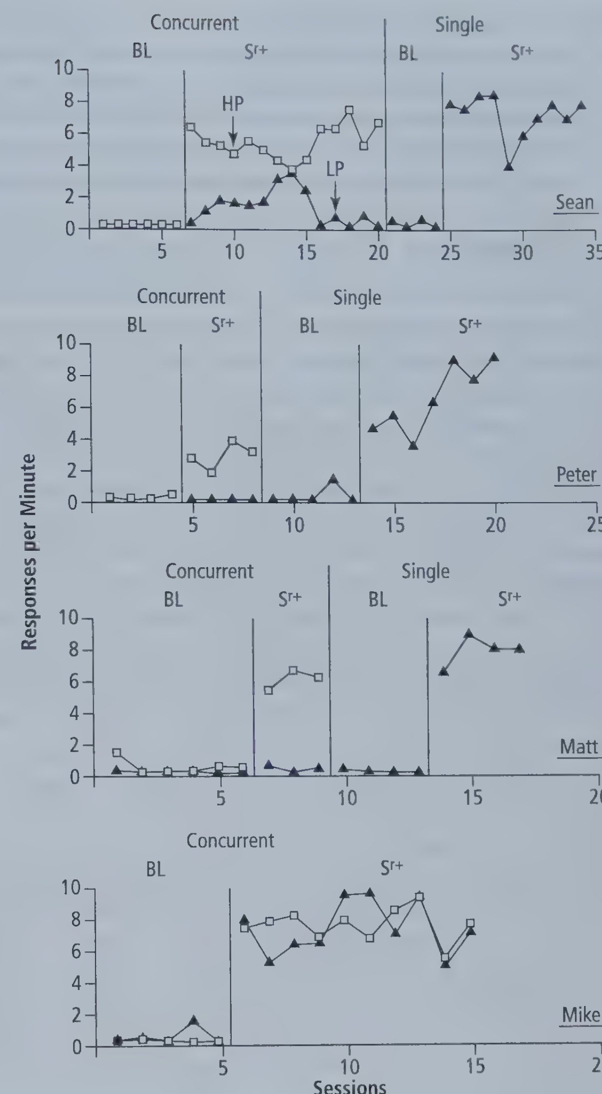


Figure 11.14 Responses per minute during concurrent-schedule and single-schedule baseline and reinforcement conditions for four adults with intellectual disabilities.

"Relative Versus Absolute Reinforcement Effects: Implications for Preference Assessments" by E. M. Roscoe, B. A. Iwata, and S. Kahng, 1999, *Journal of Applied Behavior Analysis*, *32*, p. 489. Copyright 1999 by the Society for the Experimental Analysis of Behavior, Inc.

Multiple Schedule Reinforcer Assessment

A *multiple schedule of reinforcement* consists of two or more component schedules of reinforcement for a single response with only one component schedule in effect at any given time. A discriminative stimulus (S^D) signals the presence of each component schedule, and that stimulus is present as long as the schedule is in effect. One way that a multiple schedule could be used for reinforcer assessment would be to present the same stimulus event contingent (i.e., response dependent) on each occurrence of the target behavior in one component of the multiple schedule and on a fixed-time schedule (i.e., response independent) in the other component. For example, if a practitioner wanted to use a multiple schedule to assess whether social attention functioned as a reinforcer, she would provide social attention contingent on occurrences of cooperative play when one component of the multiple schedule is in effect, and during the other component the practitioner would present the same amount and kind of social attention except on a fixed-time schedule, independent of cooperative play (i.e., noncontingent reinforcement). The teacher could apply the response-dependent schedule during the morning play period, and the response-independent schedule during the afternoon play period. If social attention functioned as reinforcement, cooperative play would likely increase over its baseline rate in the morning periods, and because of no relationship with cooperative play, attention would likely have no effect in the afternoon period. This situation follows a multiple schedule because there is one class of behavior (i.e., cooperative play), a discriminative stimulus for each contingency in effect (i.e., morning and afternoon play periods), and different conditions for reinforcement (i.e., response dependent and response independent).

Progressive-Ratio Schedule Reinforcer Assessment

Stimulus preference assessments with low response requirements (e.g., FR 1) may not predict the effectiveness of the stimulus as a reinforcer when presented with higher response requirements (e.g., on an FR 10 schedule, a student must complete 10 problems to obtain reinforcement). As DeLeon, Iwata, Goh, and Worsdell (1997) stated:

> Current assessment methods may make inaccurate predictions about reinforcer efficacy when the task used in training regimens requires either more responses or more effort

> before the delivery of reinforcement . . . for some classes of reinforcers, simultaneous increases in schedule requirements may magnify small differences in preferences that are undetected when requirements are low. In such cases, a stimulus preference assessment involving low response requirements (FR1) schedules does not accurately predict the relative potency of reinforcers under increased response requirements. (pp. 440, 446)

Progressive-ratio schedules provide a framework for assessing the relative effectiveness of a stimulus as a reinforcer as response requirements increase. In a *progressive-ratio schedule of reinforcement*, the response requirements for reinforcement are increased systematically over time independent of the participant's behavior. In a progressive-ratio schedule, the practitioner gradually requires more responses per presentation of the preferred stimulus until a breaking point is reached and the response rate declines (Roane, Lerman, & Vorndran, 2001). For example, initially each response produces reinforcement (FR 1), and then reinforcement is delivered after every 2nd response (FR 2), then perhaps after every 5th, 10th, and 20th response (FR 5, FR 10, and FR 20). At some point, a preferred stimulus may no longer function as a reinforcer (Tustin, 1994).

DeLeon and colleagues (1997) used a progressive ratio within a concurrent schedule to test the relative effectiveness of two similarly preferred stimuli (e.g., cookie and cracker) and two dissimilar stimuli (e.g., drink and balloon) as reinforcers for micro switch panel pressing for Elaine and Rick, two adults with intellectual disabilities. One panel was blue and one was yellow. The experimenters placed two reinforcers on separate plates, and put one plate behind each of the panels. Each trial (24 per session for Rick; 14 per session for Elaine) consisted of the subject pushing either one of the panels and immediately receiving the item on the plate behind that panel. During the first phase, an FR 1 schedule was used (i.e., each response produced the item on the plate). Later, the response requirement for obtaining the items was gradually increased (FR 2, FR 5, FR 10, and FR 20).

Elaine and Rick made responses that produced the two dissimilar items at roughly the same rates during the FR 1 phase (see the top two tiers in Figure 11.15). As response requirements for receiving the dissimilar stimuli increased, Elaine and Rick continued to evenly allocate responding between the two panels. However, when initially equivalent and similar reinforcers (food in FR 1) were compared under increasing schedule requirements, the differences in responses rates on the two panels revealed clear

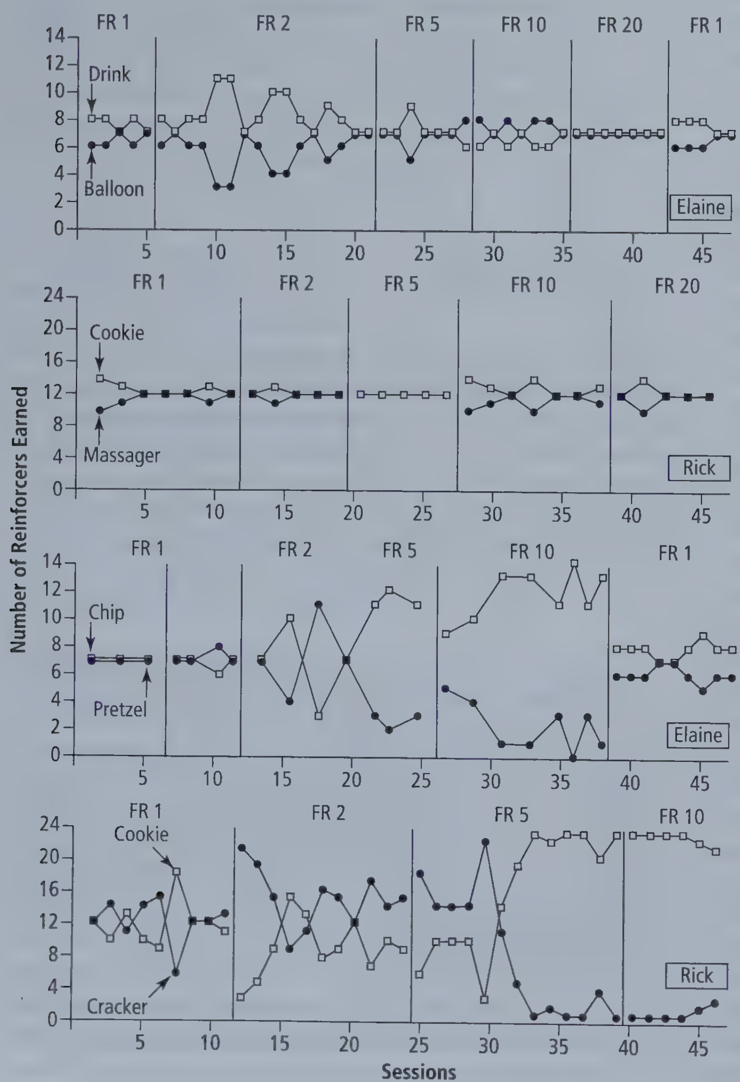


Figure 11.15 Responses per minute during concurrent-schedule and single-schedule baseline and reinforcement conditions for two adults with intellectual disabilities.

From "Emergence of Reinforcer Preference as a Function of Schedule Requirements and Stimulus Similarity," by I. G. DeLeon, B. A. Iwata, H. Goh, and A. S. Worsdell, 1997, *Journal of Applied Behavior Analysis, 30*, p. 444. Copyright 1997 by the Society for the Experimental Analysis of Behavior, Inc. Used by permission.

and consistent preferences (see the bottom two tiers in Figure 11.15). For example, when Elaine needed to work more to receive food, she allocated the majority of her responses to the panel that produced chips rather than the one that produced pretzels. In the same way, as the number of responses required to receive reinforcement increased, Rick showed a clear preference for cookies over crackers. These results suggest that "for some classes of reinforcers, simultaneous increases in schedule requirements may magnify small differences in preference that are undetected when requirements are low" (DeLeon et al., 1997, p. 446).

Increasing response requirements within a concurrent schedule may reflect the effects of increasing response requirements on the choice between reinforcers and also may reveal whether and under what conditions two reinforcers are substitutable for each other. If two reinforcers serve the same function (i.e., are made effective by the same establishing operation), an increase in the price (i.e., response requirement) for one of the reinforcers will lead to a decreased consumption of that item if a substitutable reinforcer is available (Green & Freed, 1993). DeLeon and colleagues (1997) used a hypothetical person with a slight preference for Coke over Pepsi as an analogy to explain the results shown in Figure 11.15.

> Assuming that Coke and Pepsi are both available for $1.00 per serving and that a person has only a slight preference for Coke, the individual may allocate choices rather evenly, perhaps as a function of periodic satiation for the preferred item, but with slightly more overall selections of Coke. Now assume that the cost of each is increased to $5.00 per serving. At this price, the preference for Coke is likely to be expressed. By contrast, a similar arrangement involving Coke and bus tokens may produce different results. Again, at $1.00 per item, roughly equal selection between the two options would not be surprising, assuming that the establishing operation for each dictates that both are momentarily equally valuable. However, these items serve distinctly different functions and are not substitutable; that is, the person is not free to trade one for the other and to continue to receive functionally similar reinforcement at the same rate. The person is more likely to continue choosing equally, even when the price for both reinforcers increases substantially.
>
> The same might be said for the results obtained in the present study. When choices involved two substitutable items, such as a cookie and a cracker, concurrent increases in the cost of each may have "forced" the expression of slight preference for one of the items. However, when reinforcers that were unlikely to be substitutes, such as a cookie and a massager, were concurrently available and equally preferred, increases in cost had little effect on preference. (pp. 446–447)

Although Stimulus X and Stimulus Y may each function as reinforcers when task demands are low, or when the reinforcement schedule is dense, when the task demands increase or when the schedule becomes leaner (i.e., more responses required per reinforcement), participants may choose only Stimulus Y. DeLeon and colleagues (1997) pointed out that practitioners who are alert to these relationships might be more skeptical in believing that original preferences will be sustained under changing environmental conditions, and judicious in how they plan reinforcement delivery relative to task assignment once intervention is under way. That is, it might be better to save some types of preferred stimuli for when task demands are high rather than substituting them for other equally preferred stimuli when task demands are low.

CONTROL PROCEDURES FOR POSITIVE REINFORCEMENT

Positive reinforcement control procedures are used to manipulate the contingent presentation of a potential reinforcer and observe any effects on the future frequency of behavior. *Control*, as the term is used here, requires an experimental demonstration that the presentation of a stimulus contingent on the occurrence of a target response functions as positive reinforcement. Control is demonstrated by comparing response rates in the absence and presence of a contingency, and then showing that with the absence and presence of the contingency the behavior can be turned on and off, or up and down (Baer, Wolf, & Risley, 1968). Historically, researchers and practitioners have used the reversal technique as the major control technique for positive reinforcement. Briefly, the reversal technique includes two conditions and a minimum of four phases (i.e., *ABAB*). In the *A* condition, the behavior is measured over time until it achieves stability in the absence of the reinforcement contingency. The absence of the contingency is the *control* condition. In the *B* condition, the reinforcement contingency is presented; the same target behavior continues to be measured to assess the effects of the stimulus change. The presence of the reinforcement contingency is the *experimental* condition. If the rate of responding increases in the presence of the contingency, the analyst then withdraws the reinforcement contingency and returns to the *A* and *B* conditions to learn whether the absence and presence of the contingency will turn the target behavior down and up.

However, using extinction as the control condition during the reversal phase presents practical and conceptual problems. First, withdrawing reinforcement may result in extinction-produced side effects (e.g., an initial increase in response rate, emotional responses, aggression—see Chapter 24) that affect the demonstration of control. Second, in some situations it may be impossible to withdraw the reinforcement contingency completely. For example, it is unlikely that a teacher could completely remove teacher attention during the *A* condition. In addition to these problems, Thompson and Iwata (2005) noted that, although

> extinction has often been successful in reversing the behavioral effects of positive reinforcement, its use as a control procedure presents interpretive difficulties. Essentially, extinction does not adequately isolate the reinforcement *contingency* as the variable controlling the target response, because mere stimulus presentation cannot be ruled out as an equally viable explanation. (p. 261, emphasis added)

According to Thompson and Iwata (2005), "the ideal control procedure for positive reinforcement eliminates the contingent relation between the occurrence of the target response and the presentation of the stimulus while controlling for the effects of the mere stimulus presentation" (p. 259). They reviewed the effectiveness of

three variations of the reversal technique as control procedures for determining reinforcement: noncontingent reinforcement (NCR), differential reinforcement of other behavior (DRO), and differential reinforcement of alternative behavior (DRA).[15]

Noncontingent Reinforcement

Noncontingent reinforcement (NCR) is the presentation of a potential reinforcer on a fixed-time (FT) or variable-time (VT) schedule independent of the occurrence of the target behavior. The response-independent presentation of the potential reinforcer eliminates the contingent relation between the target behavior and the stimulus presentation, while allowing any effects of the stimulus presentation alone to be detected. Thus, NCR meets Thompson and Iwata's (2005) criteria for an ideal control procedure for positive reinforcement.

The NCR reversal technique entails a minimum of five phases (*ABCBC*): *A* is a baseline condition; *B* is an NCR condition, where the potential reinforcer is presented on a fixed- or variable-interval schedule independent of the target behavior; and *C* is a condition in which the potential reinforcer is presented contingent on the occurrence of the target behavior. The *B* and *C* conditions are then repeated to learn whether the level of responding decreases and increases as a function of the absence and presence of the response-consequence contingency. The quality, amount, and rate of reinforcement should be approximately the same during the contingent and noncontingent *B* and *C* conditions of the analysis.

NCR often produces persistent responding, perhaps because of accidental reinforcement that sometimes occurs with a response-independent schedule, or because similar EOs and antecedent stimulus conditions evoke the persistent responding. Whatever the cause, persistent responding is a limitation of the NCR control procedure because it makes achieving a reversal effect (reduced responding) more time-consuming than the reversal technique with extinction. Achieving the effect may require lengthy contact with the NCR schedule.

Differential Reinforcement of Other Behavior

A practitioner using *differential reinforcement of other behavior (DRO)* delivers a potential reinforcer whenever the target behavior has not occurred during a set time interval or at a specific point in time. The DRO reversal technique includes a minimum of five phases (*ABCBC*): *A* is a baseline condition; *B* is a reinforcement condition, in which the potential reinforcer is presented contingent on the occurrence of the target behavior; and *C* is the DRO control condition in which the potential reinforcer is presented contingent on the absence of the target behavior. The analyst then repeats the *B* and *C* conditions to determine whether the level of responding decreases and increases as a function of the absence and presence of the response-consequence contingency.

The DRO schedule allows for the continued presentation of the reinforcement contingency during the reversal phases of the control procedure. In one condition, the contingency is active with occurrences of the target behavior. In another condition, the contingency is active for the omission of the target behavior. The DRO control procedure may produce the reversal effect in less time than the NCR schedule, perhaps because of the elimination of accidental reinforcement of the target behaviors.

Differential Reinforcement of Alternative Behavior

When *differential reinforcement of an alternative behavior (DRA)* is used as a control condition, the potential reinforcer is presented contingent on occurrences of a desirable alternative to the target behavior.[16] The DRA reversal technique includes a minimum of five phases (*ABCBC*): *A* is a baseline condition; *B* is a reinforcement condition, in which the potential reinforcer is presented contingent on the occurrence of the target behavior; and *C* is a condition in which the potential reinforcer is presented contingent on the occurrence of an alternative behavior (i.e., DRA). The analyst will then repeat phases *B* and *C* to ascertain whether the level of responding decreases and increases as a function of the absence and presence of the response–consequence contingency.

Thompson and Iwata (2005) summarized the limitations of using DRO and DRA as control conditions procedures to test for positive reinforcement:

> [DRO and DRA] introduce a new contingency that was not present in the original experimental arrangement. As a result, reductions in the target response under a contingency reversal might be attributed to either (a) termination of the contingency between the target response and the reinforcer or (b) introduction of reinforcement for the absence of the target response or for the occurrence of a competing response. In addition, given that reinforcement is provided contingent on some characteristic of responding during the contingency reversal, it may be difficult to control for the rate of stimulus presentation across experimental and control conditions. If responding is not quickly reduced (DRO) or reallocated toward responses that produce reinforcement (DRA), the rate of reinforcement in the control condition may be low relative to the rate of reinforcement in the experimental conditions. When this occurs, the contingency-reversal strategy is functionally similar to the conventional extinction procedure. (p. 267)

Given the considerations for the reversal technique with extinction, and its three variations, Thompson and Iwata (2005) concluded that NCR offers the most thorough and unconfounded demonstration of the effects of positive reinforcement.

USING REINFORCEMENT EFFECTIVELY

We offer practitioners nine guidelines for applying positive reinforcement effectively. These guidelines come from three main sources: the research literatures of the experimental analysis of behavior, applied behavior analysis, and our personal experiences.

Set an Easily Achieved Initial Criterion for Reinforcement

A common mistake in applications of reinforcement is setting the initial criterion for reinforcement too high, which prohibits the learner's behavior from contacting the contingency. To

use reinforcement effectively, practitioners should establish an initial criterion so that the participant's first responses produce reinforcement, and then increase the criterion for reinforcement gradually as performance improves. Heward (1980) suggested the following method for establishing initial criteria for reinforcement based on the learner's level of responding during baseline (see Figure 11.16).

> For a behavior you wish to increase, set the initial criterion higher than the child's average baseline performance and lower than or equal to his best performance during baseline. For a behavior you want to decrease in frequency, the initial criterion for reinforcement should be set below the child's average performance during baseline and greater than or equal to his lowest baseline performance. (p. 7)

Use High-Quality Reinforcers of Sufficient Magnitude

Reinforcers that maintain responding on simple tasks may not have the potency to produce similar levels of responding on more difficult or longer tasks. Practitioners will likely need to use a reinforcer of higher quality for behaviors that require more effort or endurance. A highly preferred stimulus chosen during preference assessments sometimes functions as a high-quality reinforcer. Neef, Mace, Shea, and Shade (1992), for example, found that behaviors that received a lower reinforcer rate but a higher quality reinforcer increased in occurrence, whereas behaviors that received a higher reinforcer rate with a lower quality reinforcer decreased in occurrence. Reinforcer quality is relative also to other consequences for responding currently available to the learner.

Applied behavior analysts define the magnitude (or amount) of a reinforcer as (a) the duration of time for access to the reinforcer, (b) the number of reinforcers per unit of time (i.e., *reinforcer rate*), or (c) the intensity of the reinforcer. Increases in reinforcer magnitude may correlate with an increased effectiveness of the behavior-reinforcer relation. However, the effects of reinforcer magnitude are not well understood because "few applied studies have examined the effects of magnitude on responding in a single-operant arrangement" (Lerman, Kelly, Vorndran, Kuhn, & LaRue, 2002, p. 30). Consideration of how much reinforcement to use should follow the maxim "Reinforce abundantly, but don't give away the store." We suggest that the amount of reinforcement be proportional to the quality of the reinforcer and the effort required to emit the target response.

Use Varied Reinforcers to Maintain Potent Establishing Operations

Reinforcers often decrease in effectiveness with frequent use. Presenting an overabundance of a specific reinforcer is likely to diminish the momentary effectiveness of the reinforcer due to satiation. Practitioners can minimize satiation effects by using a variety of reinforcers. If reading a specific book on sports functions as a reinforcer and the teacher relies solely on this reinforcer,

Figure 11.16 Using data from learners' baseline performance to set initial criteria for reinforcement.

The criterion-setting formulas are

For increasing behavior:

baseline average < initial criterion ≤ highest performance during baseline

For decreasing behavior:

baseline average > initial criterion ≥ lowest performance during baseline

Examples

Target Behavior	Performance Goal	Lowest	Highest	Baseline Average	Range for Initial Criterion
Playing alone	Increase	2 min.	14 min.	6 min.	7–14 min.
Identifying letters of the alphabet	Increase	4 letters	9 letters	5 letters	6–9 letters
Number of leg exercises completed	Increase	0	22	8	9–22
Percentage of math problems correctly solved	Increase	25%	60%	34%	40–60%
Number of typing errors in one letter	Decrease	16	28	22	16–21
Number of calories consumed per day	Decrease	2,260	3,980	2,950	2,260–2,900

From: "A Formula for Individualizing Initial Criteria for Reinforcement," by W. L. Heward, 1980, *Exceptional Teacher, 1*(9), p. 8. Used by permission.

ultimately reading that book may no longer produce reinforcement. Conversely, known reinforcers that are not always available may have increased effectiveness when they are reintroduced. If a teacher has demonstrated that "being first in line" is a reinforcer, but uses this reinforcer only once per week, the reinforcement effect will be greater than if "being first in line" is used frequently.

Varying reinforcers may enable less preferred stimuli to function as reinforcers. For example, Bowman, Piazza, Fisher, Hagopian, and Kogan (1997) found that some learners responded better to a variety of less preferred stimuli as compared to a continuous access to a single, more highly preferred stimulus. Also, using a variety of reinforcers may keep the potency of any one particular reinforcer higher. For example, Egel (1981) found that students' correct responding and on-task behavior were higher when the students had access to one of three randomly selected reinforcers across trials versus a constant reinforcement condition in which one of the stimuli was presented following each successful trial. Even within a session, teachers could let students select a variety of consequences from a menu. Similarly, varying a property of a reinforcer may keep its reinforcing potency for a longer time. If comic books are used as reinforcers, having several different genres of comic books available is likely to maintain their potency.

Use Direct Rather Than Indirect Reinforcement Contingencies When Possible

With a direct reinforcement contingency, emitting the target response produces direct access to the reinforcer; the contingency does not require any intervening steps. With an indirect reinforcement contingency, the response does not produce reinforcement directly. The practitioner presents the reinforcer. Research suggests that direct reinforcement contingencies may enhance performance (Koegel & Williams, 1980; Williams, Koegel, & Egel, 1981). Thompson and Iwata (2000), for example, linked the definitions of direct and indirect contingencies to the difference between automatic reinforcement (i.e., direct) and **socially mediated contingencies** (i.e., indirect) and summarized their research on response acquisition under direct and indirect contingencies of reinforcement this way:

> Under both contingencies, completion of identical tasks (opening one of several types of containers) produced access to identical reinforcers. Under the direct contingency, the reinforcer was placed inside the container to be opened; under the indirect contingency, the therapist held the reinforcer and delivered it to the participant upon task completion. One participant immediately performed the task at 100% accuracy under both contingencies. Three participants showed either more immediate or larger improvements in performance under the direct contingency. The remaining two participants showed improved performance only under the direct reinforcement contingency. Data taken on the occurrence of "irrelevant" behaviors under the indirect contingency (e.g., reaching for the reinforcer instead of performing the task) provided some evidence that these behaviors may have interfered with task performance and that their occurrence was a function of differential stimulus control. (p. 1)

Whenever possible, practitioners should use direct reinforcement contingency, especially with learners having limited behavioral repertoires.

Combine Response Prompts and Reinforcement

Response prompts are supplementary antecedent stimuli used to occasion a correct response in the presence of an S^D that will eventually control the behavior. Applied behavior analysts give response prompts before or during the performance of a target behavior. The three major forms of response prompts are instructions, modeling, and physical guidance.

Instructions describing the contingency may function as a motivating operation for learners with verbal skills, making it more likely they will contact the reinforcer more quickly. For example, Mayfield and Chase (2002) told college students who were learning five basic algebra rules that they would earn money for correct answers and not be penalized for incorrect answers.

Bourret, Vollmer, and Rapp (2004) used verbal response prompts during an assessment of the vocal verbal mand repertoires of three participants with autism.

> Each vocalization assessment session consisted of 10 trials, each 1 min in duration. A nonspecific prompt [describing the contingency] was delivered (e.g., "If you want this, ask me for it") 10 s after the onset of the trial. A prompt including a model of the complete targeted utterance (e.g., "If you want this, say 'chip' ") was delivered 20 s into the trial. The participant was prompted to say just the first phoneme of the targeted response (e.g., "If you want this, say 'ch' ") 30 s after the initiation of the trial. (pp. 131–132)

Chapter 17 provides further discussion of response prompts, including specific procedures for combining response prompts with reinforcement, and additional examples of using verbal instruction, modeling, and physical guidance response prompts.

Initially Reinforce Each Occurrence of the Behavior Initially

Provide reinforcement for each occurrence of the target behavior (i.e., continuous reinforcement) to strengthen behavior, primarily during the initial stages of learning a new behavior. After the behavior is established, gradually thin the rate of reinforcement so that some, but not all, occurrences of the behavior are reinforced (i.e., intermittent reinforcement). For example, a teacher might initially reinforce each correct response to sight words printed on flash cards and then use a ratio schedule to thin reinforcement. To firm the responses after the initial learning, provide reinforcement following two correct responses for a few trials, then following each set of four correct responses, and so on. Hanley and colleagues (2001) gradually shifted from a very dense fixed interval (FI) 1-sec schedule of reinforcement (on an FI schedule, the first target response following the end of the interval produces reinforcement) to thinner schedules with the following increments of intervals: 2 sec, 4 sec, 8 sec, 16 sec, 25 sec, 35 sec, 46 sec, and finally to an FI 58-sec schedule. Chapter 13 provides more information on using continuous and intermittent reinforcement.

Provide Contingent Attention and Descriptive Praise

As discussed earlier in this chapter, social attention and praise are powerful reinforcers for most people. However, behavioral improvements following praise often involve something more, or altogether different from, the direct effects of reinforcement. Michael (2004) discussed the common conceptual mistake of assuming that increased responding following praise and attention is a function of reinforcement.

> Consider the common use of *descriptive praise*, providing some general sign of social approval (a smile plus some comment such as "Good work!") *and, in addition*, a brief description of the behavior that is responsible for the approval ("I like the way you're . . . !"). When such praise is provided to a normally verbal person over 5 or 6 years of age, it probably functions as a form of instruction or as a rule, much as if the praiser had said, "If you want my continued approval you have to . . . " For example, a factory supervisor walks up to an employee who is cleaning up an oil spill on the factory floor, smiles broadly, and says, "George, I really like the way you're cleaning up that spill before anyone steps in it. That's very considerate of you." Now suppose that George cleans up spills from that time forward—a rather large change in behavior considering that it was followed by only a single instance of reinforcement. We might suspect that the praise functioned not simply as reinforcement but rather as a form of rule or instruction, and that George, for various reasons, provided himself with similar instruction every time another spill occurred. (pp. 164–165, emphasis in original)

A study by Goetz and Baer (1973) investigating the effects of teacher praise on preschool children's creative play with building blocks used descriptive praise in one condition of the study. "The teacher remarked with interest, enthusiasm, and delight every time that the child placed and/or rearranged the blocks so as to create a form that had not appeared previously in that session's construction(s). . . . 'Oh, that's very nice—that's different'!)" (p. 212). The three 4-year-old girls increased the construction of block form diversity during each phase of contingent descriptive praise. Goetz and Baer did not conduct a component analysis to determine how much of the girls' improved performance could be attributed to reinforcement in the form of positive attention ("That's very nice!") or to the feedback they received ("That's different!"), which enabled them to create a rule to follow ("Building *different things* with the blocks gets the teacher's attention."). The authors surmised that

> for some children, either [reinforcing attention or descriptive praise] will be sufficient without the other, but that for other children, the mix of the two will be more effective than either alone. If so, then for applied purposes a package of positive attention and descriptive praise is probably the best technique to apply to children in general. (p. 216, words in brackets added)

We recommend that, in the absence of data showing that attention and praise have produced counter-therapeutic effects for a given learner, practitioners incorporate contingent praise and attention into any intervention entailing positive reinforcement.

Gradually Increase the Response-to-Reinforcement Delay

We recommended in a previous guideline that practitioners reinforce each occurrence of the target behavior during the initial stages of learning, and then thin the delivery of reinforcers by switching to an intermittent schedule of reinforcement. Because the consequences that maintain responding in natural environments are often delayed, Stromer, McComas, and Rehfeldt (2000) reminded us that using continuous and intermittent schedules of reinforcement might be just the first steps of programming consequences for everyday situations. "Establishing the initial instances of a behavioral repertoire typically requires the use of programmed consequences that occur immediately after the target response occurs. However, the job of the applied behavior analyst also involves the strategic use of delayed reinforcement. Behaviors that yield delayed reinforcement are highly adaptive in everyday life, but they may be difficult to establish and maintain" (p. 359).[17]

Examples of tactics that applied behavior analysts have used to help people learn to respond effectively for delayed consequences include (a) a delay-to-reinforcement time interval that begins with a short delay that is gradually increased (Dixon, Rehfeldt, & Randich, 2003; Schweitzer & Sulzer-Azaroff, 1988); (b) a gradual increase in work requirements during the delay (Dixon & Holcomb, 2000); (c) an activity during the delay to "bridge the gap" between the behavior and reinforcer (Mischel, Ebbesen, & Zeiss, 1972); and, importantly, (d) verbal instruction in the form of an assurance that the reinforcer will be available following a delay (e.g., "The calculator will show the amount of money to be placed in a savings account for you. You will be given all the nickels in your savings account on [day]") (Neef, Mace, & Shade, 1993, p. 39). We present more on using delayed consequences to promote generalization and maintenance of behavior changes in Chapter 30.

Gradually Shift from Contrived to Naturally Occurring Reinforcers

We end this chapter with an extract from Murray Sidman's (2000) insightful and thought-provoking account of what he learned in the "early days" of applying behavioral principles to human behavior. In describing a project from 1965 to 1975 that emphasized the use of positive reinforcement with boys between the ages of 6 and 20 years who were diagnosed with intellectual disabilities and living in a state institution, Sidman recollected how introducing tokens as generalized conditioned reinforcers eventually led to praise from the project staff, and later to learning itself, becoming powerful reinforcers for the boys.

> We began with tokens, which had the advantage of being visible and easily handled. Later, after the boys had learned to save tokens and to understand numbers, we were able to introduce points. For some, points led eventually to money. As the boys saw how pleased we were when they earned the tokens and points that brought them other reinforcers,

our pleasure also became important to them, and we became able to use praise as a reinforcer. As they learned more and more, many of the boys found that what they learned permitted them to deal more effectively with their gradually enlarging world. For them, learning itself became reinforcing. (p. 19)

Success in manipulating the environment may be the ultimate naturally occurring reinforcer. As Skinner (1989) pointed out, this powerful reinforcer "does not need to be contrived for instructional purposes; it is unrelated to any particular kind of behavior and hence always available. We call it *success*." (p. 91)

SUMMARY

Positive Reinforcement Defined

1. Positive reinforcement occurs when a response is followed immediately by the presentation of a stimulus change that increases the future occurrence of similar responses.
2. The stimulus that is presented as a consequence, and that is responsible for the subsequent increase in responding, is called a positive reinforcer, or, more simply, a reinforcer.
3. The importance of the immediacy of reinforcement must be emphasized; a response-to-reinforcement delay of as little as 1 second will be less effective than a reinforcer delivered immediately.
4. The effects of long-delayed consequences on human behavior should not be attributed to reinforcement, as these effects drop off quickly as the delay is increased.
5. A misconception held by some is that reinforcement is a circular concept. Circular reasoning is a form of faulty logic in which cause and effect are confused and not independent of each other. Reinforcement is not a circular concept because the two components of the response–consequence relation can be separated and the consequence manipulated to determine whether it increases the frequency of the behavior it follows.
6. Reinforcement does more than increase the future occurrence of similar behavior; it also changes the function of stimuli that immediately precede the reinforced behavior. By virtue of being temporally paired with the response–reinforcer contingency, antecedent events acquire the ability to evoke (make more likely) instances of the reinforced response class. An antecedent stimulus that evokes behavior because it has been correlated with the availability of reinforcement is called a discriminative stimulus (S^D).
7. A discriminated operant is defined by a three-term contingency of $S^D \rightarrow R \rightarrow S^{R+}$.
8. The momentary effectiveness of any stimulus change as reinforcement depends on an existing level of motivation with respect to that stimulus change. An establishing operation (EO) (e.g., deprivation) increases the current effectiveness of a reinforcer; an abolishing operation (AO) (e.g., satiation) decreases the current effectiveness of a reinforcer.
9. A complete description of reinforcement of a discriminated operant entails a four-term contingency: $EO \rightarrow S^D \rightarrow R \rightarrow S^{R+}$.
10. Automaticity of reinforcement refers to the fact that a person does not have to understand or verbalize the relation between his or her actions and a reinforcing consequence, or for that matter even be aware that a consequence has occurred, for reinforcement to happen.
11. Reinforcement strengthens any behavior that immediately precedes it; no logical or adaptive connection between behavior and the reinforcing consequence is necessary.
12. The importance of understanding the arbitrariness of reinforcement goes far beyond providing a possible explanation for the development of harmless superstitious and idiosyncratic behaviors. The arbitrary nature of selection by reinforcement may explain the acquisition and maintenance of many maladaptive and challenging behaviors.
13. Automatic reinforcement refers to the behavior–stimulus change relation that occurs without the presentation of consequences by other people, or is assumed when a behavior persists in the absence of any known reinforcer.

Classifying Reinforcers

14. An unconditioned reinforcer is a stimulus change that functions as reinforcement even though the learner has had no particular learning history with it. (The terms *primary reinforcer* and *unlearned reinforcer* are synonyms for unconditioned reinforcer.)
15. A conditioned reinforcer (sometimes called a *secondary reinforcer* or *learned reinforcer*) is a previously neutral stimulus change that has acquired the capability to function as a reinforcer through stimulus–stimulus pairing with one or more unconditioned reinforcers or conditioned reinforcers.
16. A generalized conditioned reinforcer is a conditioned reinforcer that, as a result of having been paired with many unconditioned and conditioned reinforcers, does not depend on a current EO for any particular form of reinforcement for its effectiveness.
17. When reinforcers are described by their physical properties, they are typically classified as edible, sensory, tangible, activity, or social reinforcers.
18. The Premack principle states that making the opportunity to engage in a behavior that occurs at a relatively high free operant (or baseline) rate contingent on the occurrence of low-frequency behavior will function as reinforcement for the low-occurrence behavior.
19. The response-deprivation hypothesis is a model for predicting whether access to one behavior (the contingent behavior) will function as reinforcement for another

behavior (the instrumental response) based on the relative baseline rates at which each behavior occurs and whether access to the contingent behavior represents a restriction compared to the baseline level of engagement.

Identifying Potential Reinforcers

20. Stimulus preference assessment refers to a variety of procedures used to determine (a) the stimuli that the person differentially selects; (b) the relative hierarchical preference value of those stimuli (high preference to low preference); (c) the conditions under which those preference values change when task demands, deprivation states, or schedules of reinforcement are modified; and (d) whether highly preferred items ultimately serve as effective reinforcers.

21. Stimulus preference assessment can be conducted using at least five forms: the single-stimulus (SS), paired-stimulus (PS), free operant (FO), multiple-stimulus with item replacement (MSWI), multiple-stimulus without item replacement (MSWO). These five forms can be clustered into three categories: ask the target person and/or significant others, conduct free operant observations, and conduct trial-based assessments.

22. Preferred stimuli do not always function as reinforcers, and stimulus preferences often change over time.

23. Reinforcer assessment refers to a variety of direct, data-based methods for determining the relative effects of a given stimulus as reinforcement under different and changing conditions or the comparative effectiveness of multiple stimuli as reinforcers for a given behavior under specific conditions. Reinforcer assessment is often conducted with in-the-moment procedures, concurrent schedules of reinforcement, multiple schedules of reinforcement, and progressive reinforcement schedules.

Control Procedures for Positive Reinforcement

24. Positive reinforcement control procedures are used to manipulate the contingent presentation of a potential reinforcer and observe any effects on the future frequency of behavior. *Control*, as the term is used here, requires an experimental demonstration that the presentation of a stimulus contingent on the occurrence of a target response functions as positive reinforcement. Control is demonstrated by comparing response rates in the absence and presence of a contingency, and then showing that with the absence and presence of the contingency the behavior can be turned on and off, or up and down.

25. In addition to a reversal design using the withdrawal of the reinforcement contingency (i.e., extinction) as the control condition, noncontingent reinforcement (NCR), differential reinforcement of other behavior (DRO), and differential reinforcement of alternative behavior (DRA) can be used as control conditions for reinforcement.

Using Reinforcement Effectively

26. Guidelines for increasing the effectiveness of positive reinforcement interventions include:

- Set an easily achieved initial criterion for reinforcement.
- Use high-quality reinforcers of sufficient magnitude.
- Use varied reinforcers to maintain potent establishing operations.
- Use direct rather than indirect reinforcement contingencies when possible.
- Combine response prompts and reinforcement.
- Initially reinforce each occurrence of the behavior.
- Provide contingent attention and descriptive praise.
- Gradually increase the response-to-reinforcement delay.
- Gradually shift from contrived to naturally occurring reinforcers.

KEY TERMS

automatic reinforcement
conditioned reinforcer
generalized conditioned reinforcer
positive reinforcement
positive reinforcer
Premack principle
reinforcer assessment
response-deprivation hypothesis
rule-governed behavior
socially mediated contingencies
stimulus preference assessment
unconditioned reinforcer

MULTIPLE-CHOICE QUESTIONS

1. A teacher wants to deliver a sticker to Johnny every time he says "Please." This would be considered positive reinforcement if:

a. The behavior of saying "Please" increases over time.
b. The behavior of saying "Please" decreases over time.
c. The behavior of saying "Please" does not change over time.
d. The behavior of sticker giving decreases over time.

Hint: (See "Positive Reinforcement Defined")

2. The qualifications to be considered when utilizing positive reinforcement are:

a. Delay between response and consequence
b. Stimulus conditions when response was emitted
c. Strength of motivation
d. All of these

Hint: (See "Using Reinforcement Effectively")

3. An example of unconditioned reinforcer(s) is:
   a. Food
   b. Water
   c. Oxygen
   d. Warmth
   e. Sexual stimulation
   f. All of these

   Hint: (See "Classifying Reinforcers")
4. A neutral stimulus can be made into a reinforcer by:
   a. Pairing it with an unconditioned stimulus
   b. Pairing it with an unconditioned response
   c. Presenting it individually after a response multiple times
   d. A neutral stimulus cannot be utilized as a reinforcer

   Hint: (See "Classifying Reinforcers")
5. Heather hears a phone ring, then picks up the phone. Her husband talks with her over the phone. What is the behavior in this three-term contingency?
   a. Husband talking
   b. Heather picks up the phone
   c. Phone ringing
   d. None of these

   Hint: (See "Positive Reinforcement Defined")
6. An example of a motivating operation is:
   a. Hunger
   b. Hay fever
   c. Argument with significant other
   d. All of these
   e. None of these

   Hint: (See "Positive Reinforcement Defined")
7. Which of the following is not an appropriate way to identify potential reinforcers?
   a. Pick a stimulus that is reinforcing for someone else
   b. Ask the person or significant other what is reinforcing
   c. Observe the person
   d. Measure the person's responses to trial-based tests

   Hint: (See "Identifying Potential Reinforcers")
8. Stimulus preference assessment refers to a variety of procedures to determine:
   a. The stimulus a person prefers
   b. The relative preference values of those stimuli
   c. The conditions under which those preference values hold
   d. All of these

   Hint: (See "Identifying Potential Reinforcers")
9. What information does a concurrent schedule reinforcer assessment provide?
   a. A rank order of preference for a variety of stimuli
   b. How often children will select one stimulus versus another to indicate preference
   c. Whether or not a stimulus functions as a reinforcer and how effective that stimulus is as a reinforcer compared to other stimuli
   d. All of these
   e. None of these

   Hint: (See "Reinforcer Assessment")
10. When two or more contingencies of reinforcement operate independently and simultaneously for two or more behaviors, what schedule of reinforcement is in effect?
    a. Multiple
    b. Mixed
    c. Progressive-Ratio
    d. Concurrent
    e. All of these
    f. None of these

    Hint: (See "Reinforcer Assessment")
11. Which schedule reinforcer assessment consists of two or more component schedules of reinforcement for a single response with only one component schedule in effect at any given time?
    a. Progressive-Ratio
    b. Multiple
    c. Mixed
    d. Concurrent
    e. All of these
    f. None of these

    Hint: (See "Reinforcer Assessment")
12. The model for predicting whether (a) access to one behavior will function as reinforcement for another behavior based on the relative baseline rates at which each behavior occurs and (b) whether contingent access to one of the behaviors represents a restriction compared to the baseline level of engagement is known as the:
    a. Response-satiation hypothesis
    b. Response-contingency hypothesis
    c. The Premack Principle
    d. Response-deprivation hypothesis
    e. None of these

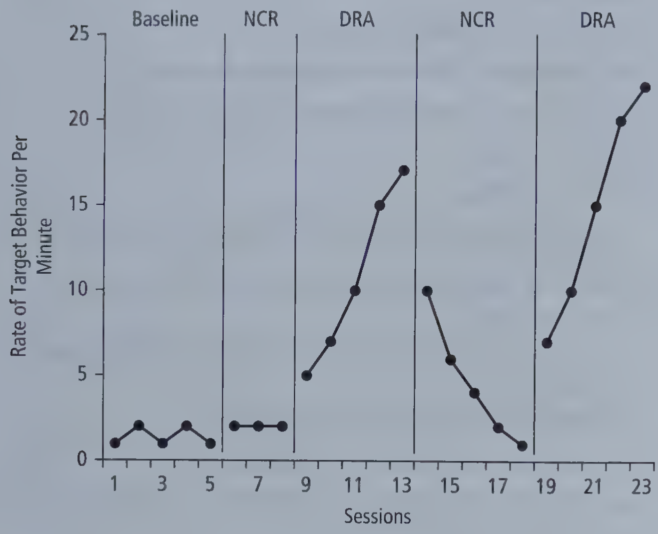
    Hint: (See "Classification of Reinforcers by Formal Properties")

13. The response-deprivation hypothesis was built on whose concept?
   a. Premack
   b. Skinner
   c. Watson
   d. Pavlov
   e. None of these
   Hint: (See "Classification of Reinforcers by Formal Properties")
14. Experimental control when evaluating positive reinforcement is demonstrated by comparing:
   a. Response rates in the presence of a contingency
   b. Response rates in the absence of a contingency
   c. Fixed time reinforcement to differential reinforcement
   d. Differential reinforcement of other behavior to another reinforcement schedule, such as differential reinforcement of alternate behavior
   e. All of these
   f. None of these
   Hint: (See "Control Procedures for Positive Reinforcement")
15. The presentation of a potential reinforcer on a fixed or variable schedule independent of the occurrence of the target behavior is also referred to as:
   a. Contingent reinforcement
   b. Noncontingent reinforcement
   c. Contingent punishment
   d. Noncontingent punishment
   e. None of these
   f. All of these
   Hint: (See "Control Procedures for Positive Reinforcement")
16. Which of the following could be used as a control condition when evaluating the effects of a reinforcement-based procedure?
   a. Noncontingent reinforcement
   b. Differential reinforcement of other behavior
   c. Differential reinforcement of alternative behavior
   d. All of these
   e. None of these
   Hint: (See "Control Procedures for Positive Reinforcement")
17. Which of the following are guidelines for effective reinforcement?
   a. Choosing reinforcers that are not relevant to the target behavior
   b. Using reinforcers of low magnitude
   c. Delaying the delivery of a reinforcer
   d. Using the same reinforcers over and over again
   e. None of these
   Hint: (See "Using Reinforcement Effectively")
18. Practitioners increase the effectiveness of positive reinforcement when they attend to the ________ of stimulus presentation.
   a. Arrangement
   b. Management
   c. Administration
   d. All of these
   e. None of these
   Hint: (See "Using Reinforcement Effectively")

ESSAY-TYPE QUESTIONS

1. Briefly explain the terms positive reinforcement and positive reinforcer.
   Hint: (See "Positive Reinforcement Defined")
2. Briefly explain the terms conditioned and unconditioned reinforcers and give an example of each.
   Hint: (See "Classifying Reinforcers")
3. Bob has just completed his morning workout routine. On the way to taking a shower, Bob walks past the refrigerator. Bob takes an apple from the refrigerator. Bob eats the apple on the way to taking his morning shower. Briefly describe the above example using the operant conditioning paradigm. (Be sure to include motivation in your answer.)
   Hint: (See "Positive Reinforcement Defined")
4. Briefly describe both the pros and cons of the three basic methods for stimulus preference assessments. (The three methods are asking the person or significant others, observing the individual, and measuring the person's responses to trial-based tests.)
   Hint: (See "Identifying Potential Reinforcers")
5. After identifying possible reinforcers how could it be determined if the stimulus is a reinforcer? Briefly describe what assessment procedure you would use and why.
   Hint: (See "Identifying Potential Reinforcers")
6. Describe and give an example of the response-deprivation procedure. Explain your answer.
   Hint: (See "Classifying Reinforcers")
7. Briefly explain whether or not "control" has been demonstrated in the graph. Support your explanation with evidence from the graph. (Control requires an experimental demonstration that the presentation of a stimulus

contingent on the occurrence of a target response functions as a positive reinforcement.)



Hint: (See "Control Procedures for Positive Reinforcement")

8. Shawn is overweight and does not currently do any form of daily exercise. His doctor has recommended that he walk 3 miles daily. Briefly explain how positive reinforcement could be utilized to help Shawn increase the behavior of walking. (Include a proposed reinforcer, timing of proposed reinforcer delivery, contingency, motivation, and a statement about the proposed way of evaluating the effectiveness of the proposed reinforcer).

   Hint: (See "Using Reinforcement Effectively")

NOTES

1. Phrases such as *strengthening the behavior* and *increasing the likelihood of future responding* are sometimes used to describe the basic effect of reinforcement. Although such terms appear occasionally in this book, recognizing Michael's (1995) concern that use of such terms "encourages a language of intervening variables, or an implied reference to something other than an observable aspect of behavior" (p. 274), we most often use *increased future occurrence (or rate)* to refer to the primary effect of reinforcement.
2. When the consequence that produced the increase in responding is best described as the *termination* or *withdrawal* of an already present stimulus, *negative reinforcement* has occurred. The fundamental nature and qualifying conditions for positive reinforcement and negative reinforcement are the same: Behavior increases in the future. Negative reinforcement is examined in detail in Chapter 12.
3. Excellent discussions of rule-governed behavior can be found in Baum (2017); Chase and Danforth (1991); Hayes (1989); Hayes, Zettle, and Rosenfarb (1989); Malott and Shane (2014); Reitman and Gross (1996); Schlinger and Blakely (1987); and Vaughan (1989).
4. Motivating operations are described in detail in Chapter 16.
5. It is a mistake to assume that all superstitious behavior is the direct result of adventitious reinforcement. Many superstitious behaviors are probably the result of following cultural practices. For example, high school baseball players may wear their caps inside out and backwards when a late-inning rally is needed because they have seen major leaguers don such "rally caps" in the same situation.
6. Mands, tacts, and intraverbals—three elementary verbal operants first described by Skinner (1957)—are explained in Chapter 18.
7. Remember, the environment does the pairing, not the learner. The learner does not have to "associate" the two stimuli.
8. The first volume of the *Journal of Applied Behavior Analysis* (1968) is a treasure trove of classic studies in which simple and elegant experimental designs revealed the powerful effects of operant conditioning and contingency management. We strongly encourage any serious student of applied behavior analysis to read it from cover to cover, and to return to it now and then for inspiration.
9. See https://www.kennedykrieger.org/sites/default/files/patient-care-files/paired_stimulus_preference_assessment.pdf for detailed procedural steps and associated sample forms to complete a variety of SPAs.
10. For additional details on conducting single-stimulus assessments, including a sample data sheet, refer to https://www.kennedykrieger.org/sites/default/files/patient-care-files/single_stimulus_preference_assessment.pdf and Hagopian et al. (2001).
11. For additional details on conducting paired-stimulus preference assessments, including a sample data sheet, refer to https://www.kennedykrieger.org/sites/default/files/patient-care-files/paired_stimulus_preference_assessment.pdf and Fisher et al. (1992).
12. For additional details on conducting multiple stimulus without replacement (MSWO) preference assessments, including a sample data sheet, refer to https://www.kennedykrieger.org/sites/default/files/patient-care-files/mswo_preference_assessment.pdf and DeLeon and Iwata (1996).
13. For additional information on comparative effects of SPA, see Verriden and Roscoe (2016) and Karsten et al. (2011).
14. These schedules of reinforcement and their effects on behavior are described in Chapter 13.
15. Chapter 8 presents the ABAB, NCR, DRO, and DRA control techniques in the context of single-case experimental designs, and Chapter 25 discusses NCR, DRO, and DRA in detail.
16. Chapter 25 describes DRO and DRA as behavior change tactics for decreasing the frequency of undesirable behavior.
17. Moving from continuous reinforcement to an intermittent schedule of reinforcement is sometimes described as a means of increasing reinforcer delay (e.g., Alberto & Troutman, 2013; Kazdin, 2013). However, an intermittent schedule of reinforcement does not entail "delayed reinforcement" unless specified. Although only some occurrences of the target behavior are reinforced on an intermittent schedule of reinforcement (see Chapter 13), reinforcement is delivered immediately following the response that meets the contingency. For example, on a fixed ratio 10 schedule of reinforcement, every 10th response produces immediate reinforcement. *Delay to reinforcement* or *reinforcement delay* describes the time lapse between the response and delivery of the reinforcer after the contingency has been met (e.g., the reinforcer was delivered 45 seconds after every 10th response).

CHAPTER 12

Negative Reinforcement

Richard G. Smith and Brian A. Iwata

LEARNING OBJECTIVES

- Define and provide examples of negative reinforcement.
- Identify and use negative reinforcers.
- Differentiate between escape and avoidance contingencies.
- Identify the characteristics of negative reinforcement.
- Use appropriate parameters and schedules of negative reinforcement.
- State and plan for the possible unwanted effects of and ethical issues in the use of negative reinforcement.

Have you ever been "that person" whose cell phone rings in an otherwise quiet classroom? If so, you probably felt the burning stares of the professor and the other students in the class as you quickly scrambled to turn off your ringer and sank deeply into your chair. You probably also make sure to turn your ringer off before you attend classes now. These are both instances of negative reinforcement contingencies, in which behavior increases when it results in escape from or avoidance of aversive events. Turning off the ringer in class produced escape from the judgmental stares of your classmates, and turning off the ringer before class now allows you to avoid any such future embarrassment.

As noted in Chapter 11, positive reinforcement involves an increase in responding as a function of contingent *presentation* of a stimulus. In a complementary way, responding can lead to the *termination or avoidance* of a stimulus, and when responding increases as a result of this arrangement, learning has occurred through *negative reinforcement*. This chapter expands the discussion of operant contingencies to include negative reinforcement. We define negative reinforcement, distinguish between escape and avoidance contingencies, describe events that may serve as the basis for negative reinforcement, illustrate ways negative reinforcement may be used to strengthen behavior, and discuss ethical issues that arise when using negative reinforcement. Readers interested in more in-depth discussions of basic and applied research on negative reinforcement are referred to reviews by Hineline and Rosales-Ruiz (2013) and Iwata (1987).

DEFINITION OF NEGATIVE REINFORCEMENT

A **negative reinforcement** contingency is one in which the occurrence of a response produces the termination, reduction, postponement, or avoidance of a stimulus, which leads to an increase in the future occurrence of that response. A full description of negative reinforcement requires specification of its four-term contingency (see Figure 12.1): (a) The establishing operation (EO) for behavior maintained by negative reinforcement is an antecedent event in whose presence escape (termination of the event) is reinforcing, (b) the discriminative stimulus (S^D) is another antecedent event in whose presence a response is more likely to be reinforced, (c) the response is the act that produces reinforcement, and (d) the reinforcer is the termination of the event that served as the EO.

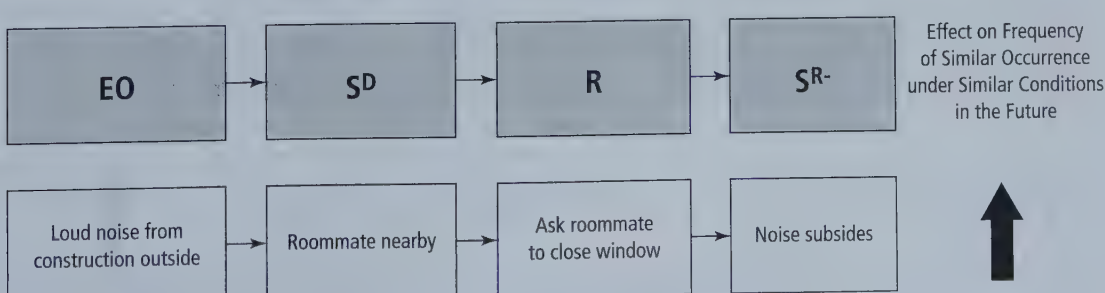


Figure 12.1 Four-term contingency illustrating negative reinforcement.

Positive Versus Negative Reinforcement

Positive and negative reinforcement have a similar effect on behavior in that both produce an increase in responding. They differ, however, with respect to the type of stimulus change that follows behavior, as illustrated in Figure 12.2. In both examples, a stimulus change (consequence) strengthens the behavior that preceded it: Asking the sibling to make a sandwich is strengthened by obtaining food; carrying rain protection is strengthened by blocking the rain. However, behavior maintained by positive reinforcement produces a stimulus that was absent prior to responding, whereas behavior maintained by negative reinforcement terminates a stimulus that was present prior to responding: Food was unavailable prior to asking for it but available after (positive reinforcement); rain was landing on one's clothing before raising the umbrella but not after (negative reinforcement).

Thus, the key distinction between positive and negative reinforcement is based on the type of stimulus change that occurs following a response. Many stimulus changes have discrete onsets and offsets and involve an "all-or-none" operation, as in turning on a television (positive reinforcement) or turning off the light in a bedroom (negative reinforcement). Other stimulus changes exist on a continuum from less to more, such as turning up the volume of a stereo to hear it better (positive reinforcement) or turning it down when it is too loud (negative reinforcement). Sometimes, however, it is difficult to determine whether an increase in responding resulted from positive or negative reinforcement because the stimulus change is ambiguous. For example, although a change in temperature can be measured quantitatively so we know that it either increased or decreased following behavior, it is unclear whether turning on a heater when the temperature is 40° F (5° C) is an example of positive reinforcement because the response "produced heat" or negative reinforcement because the response "removed cold." Another example can be found in a classic study by Osborne (1969) on the use of free time as a reinforcer in the classroom. During baseline, students were observed to get out of their seats frequently during long work periods. During treatment, the students were given 5 minutes of free time if they remained in their seats during 10-minute work periods, and in-seat behavior increased. At first glance, the free-time contingency appears to have involved negative reinforcement (termination of the in-seat requirement contingent on appropriate behavior). As Osborne noted, however, access to activities (games, social interaction, etc.) during free time may have functioned as positive reinforcement.

Given the ambiguous nature of some stimulus changes, Michael (1975) suggested that the distinction between positive and negative reinforcement, based on whether a stimulus is presented or removed, may be unnecessary. Instead, he emphasized the importance of specifying the type of environmental change produced by a response in terms of key stimulus features that constituted both the "prechange" and "postchange" conditions. This practice, he proposed, would eliminate the necessity of describing the transition between prechange and postchange conditions as one involving the presentation or removal of stimuli and would facilitate a more complete understanding of functional relations between environment and behavior.

Little has changed since the publication of Michael's (1975) article; the distinction between positive and negative reinforcement continues to be emphasized in every text on learning principles, and citations to the term *negative reinforcement* have increased in applied research (Iwata, 2006). In an attempt to renew the discussion, Baron and Galizio (2005) reiterated Michael's position and included some additional points of emphasis. This terminological issue is a complex one that can be considered from several perspectives—conceptual, procedural, and historical—and debate on the issue continues at this time. Readers interested in the topic are referred to a series of reactions to Baron and Galizio (Chase, 2006; Iwata, 2006; Lattal & Lattal, 2006;

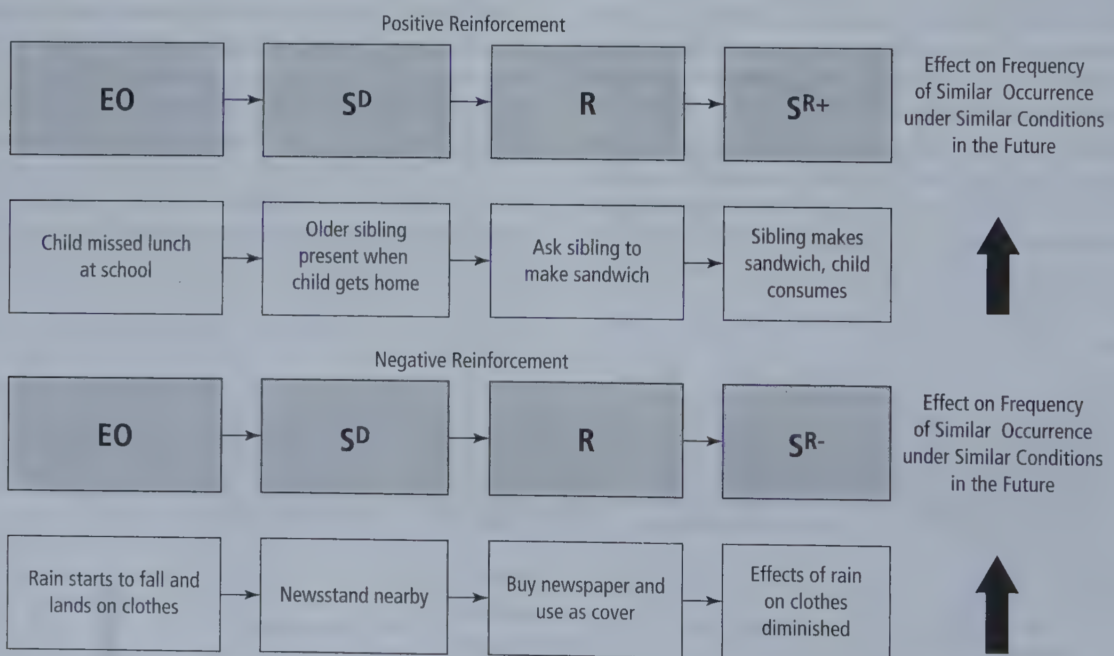


Figure 12.2 Four-term contingency illustrating similarities and the difference between positive and negative reinforcement.

Marr, 2006; Michael, 2006; Sidman, 2006) and their rejoinder (Baron & Galizio, 2006). One difficulty that arises when attempting to examine the issue empirically is equating the two types of contingencies. Magoon and colleagues did so in two human operant studies (Magoon & Critchfield, 2008; Magoon, Critchfield, Merrill, Newland, & Schneider, 2017) by pitting money-gain positive reinforcement against money-loss avoidance negative reinforcement. Their results showed differential outcomes under the two arrangements, supporting the proposition that the two types of contingencies are functionally unique.

Negative Reinforcement Versus Punishment

Negative reinforcement is sometimes confused with punishment for two reasons. First, because the lay term for positive reinforcement is *reward*, people mistakenly consider negative reinforcement as the technical term for the opposite of reinforcement (punishment). The terms *positive* and *negative*, however, do not refer to "good" and "bad" but to the type of stimulus change (presentation versus termination) that follows behavior (Catania, 2013). A second source of confusion stems from the fact that the stimuli involved in both negative reinforcement and punishment are, by definition, "aversive."[1] Although it is true that the same stimulus may serve as a negative reinforcer in one context and as a punisher in a different context, both the nature of the stimulus change and its effect on behavior differ. In a negative reinforcement contingency, a stimulus that was present is terminated by a response, which leads to an increase in responding; in a punishment contingency, a stimulus that was absent is presented following a response, which leads to a decrease in responding. Thus, a response that terminates loud noise would increase as a function of negative reinforcement, but one that produces loud noise would decrease as a function of punishment (see Chapter 14 for a more extensive discussion of punishment).

ESCAPE AND AVOIDANCE CONTINGENCIES

In its simplest form, negative reinforcement involves an **escape contingency**, in which a response terminates (produces escape from) an ongoing stimulus. An early study by Keller (1941) illustrates typical laboratory research on escape. When a rat was placed in an experimental chamber and a bright light was turned on, the rat quickly learned to press a lever, which turned off the light. Osborne's (1969) study on free-time contingencies may also serve as an example of escape learning in an applied context. To the extent that the important feature of the contingency was the termination of work requirements, in-seat behavior during 10-minute work periods produced 5 minutes of escape.

Although situations involving escape are commonly encountered in everyday life (e.g., we turn off loud noises, shield our eyes from the sun, flee from an aggressor), much behavior maintained by negative reinforcement is characterized by an **avoidance contingency**, in which a response prevents or postpones the presentation of a stimulus. Returning to the previous laboratory example, an experimenter can add to the escape contingency an arrangement in which another stimulus such as a tone precedes the presentation of the bright light, and a response in the presence of the tone eliminates the presentation of the light or postpones it until the tone is next presented. This type of arrangement has been called **discriminated avoidance**, in which responding in the presence of a signal prevents the onset of a stimulus from which escape is a reinforcer. Because responses in the presence of the tone are reinforced, whereas those in the absence of the tone have no effect, the tone is a discriminative stimulus (S^D) in whose presence there is an increased likelihood of reinforcement for responding. (See Chapter 17 for more on stimulus control.)

Avoidance behavior also can be acquired in the absence of a signal. Suppose the experimenter arranges a schedule in which the bright light turns on for 5 seconds every 30 seconds, and a response (or some number of responses) at any time during the interval resets the clock to zero. This type of arrangement is known as **free-operant avoidance** because the avoidance behavior is "free to occur" at any time and will delay the presentation of the bright light.

Each of the three types of contingencies described earlier was illustrated in an ingenious study by Azrin, Rubin, O'Brien, Ayllon, and Roll (1968) on postural slouching (see Figure 12.3). Participants wore an apparatus that closed an electrical circuit when slouching occurred. Closure of the switch produced an audible click, which was followed 3 seconds later by a 55-db

Figure 12.3 Three types of negative reinforcement contingencies used by Azrin and colleagues (1968) to maintain correct posture.

Free-Operant Avoidance

Correct posture maintained → Avoid click and tone

Discriminated Avoidance

Slouching (incorrect posture) → Audible click
Correct posture within 3 seconds of click → Avoid tone

Escape

Slouching (incorrect posture) → Audible click
Posture not corrected within 3 seconds → 55-db tone
Posture corrected → Tone turns off

tone. Postural correction in the presence of the tone turned off the tone (escape) but prevented the tone if correction occurred during the 3 seconds following the click (discriminated avoidance). Furthermore, maintenance of correct posture prevented the click (free-operant avoidance). A hypothetical example involving homework management also illustrates these contingencies. A parent who sends a child to his or her room immediately following school and allows the child to leave the room only after completing homework has arranged an escape contingency: Homework completion produces escape from the bedroom. A parent who first delivers a warning (e.g., "If you don't start your homework in 10 minutes, you'll have to do it in your bedroom") has arranged a discriminated avoidance contingency: Starting homework following the warning avoids having to do it in the bedroom. Finally, the parent who waits until later in the evening to impose the in-room requirement has arranged a free-operant avoidance contingency: Homework completion at any time after school avoids having to do it in the bedroom later.

As can be seen above, escape and avoidance contingencies illustrate very different arrangements, and one particular feature of some avoidance contingencies has raised questions about fundamental learning processes. Whereas the environment–behavior relationship in escape contingencies is clearly identifiable—that is, responding results in termination or reduction of a stimulus that was present initially—the evocative and maintaining influences involved in avoidance are less clear. Except in the case of discriminated avoidance, avoidance behavior often occurs in the absence of any apparent evocative condition and does not produce obvious consequential changes in the environment. This seems inconsistent with our A-B-C notion of operant behavior: We see the "B," but where are the "A" and "C"? How do we account for behavior that does not appear to fit a conceptual framework that emphasizes the importance of immediate antecedent and consequent events as determinants of behavior? (See Box 12.1 for further discussion of avoidance responding.)

BOX 12.1

Is Avoidance a Different Type of Learning?

Two-Factor Theory

In the 1940s, Mowrer conducted research showing that, after several trials in which a tone was presented just before a shock, rats began to emit escape responses (moving over a barrier to another location) in the presence of only the tone. Mowrer proposed that this behavior resulted from two independent processes. First, the shock was considered to be an unconditioned stimulus, eliciting reflexive "fear" responses. After being paired with the shock, the tone acquired fear-eliciting properties through respondent conditioning. Thus, Mowrer asserted that responding in the presence of the tone produced escape from the fear and anxiety elicited by the tone, rather than escape from the shock. He proposed a "two-factor" theory, incorporating both respondent and operant mechanisms, to account for avoidance behavior. Some subsequent research provided further evidence in support of this theory, such as a study by Kamin (1957) showing that avoidance behavior was disrupted when it canceled upcoming shocks but did not terminate the warning tone. Results of other research, however, have been inconsistent with Mowrer's theory. For example, Solomon, Kamin, and Wynne (1953) showed that dogs continued to emit avoidance responses in the presence of a warning signal long after shocks were no longer programmed to follow the signal. And work by Sidman (1953, 1962) and others (Herrnstein & Hineline, 1966; Perone & Galizio, 1987) showed that when shocks were presented according to time-based schedules, animals learned to postpone shocks in the absence of a programmed signal, further calling Mower's two-factor theory into question.

Anxiety as derived relational responding

Behavior analysts also have extended findings of research on derived relational responding to investigate clinical problems involving avoidance (Dymond, 2009). It has been suggested that anxiety and phobic disorders develop when aversive properties spread across classes of stimuli whose function is established through associative, or indirect, conditioning processes. For example, Dougher, Hamilton, Fink, and Harrington (2007) showed that fear responses spread in an orderly, hierarchical way across three otherwise arbitrary stimuli to which college students had been trained to respond according to a small, medium, and large relation. After the "medium" stimulus was paired with a mild shock, a smaller fear response was recorded for the "small" stimulus, whereas a larger fear response was recorded in the presence of the "large" stimulus—even though neither of these stimuli had ever been paired with the shock! The authors suggested that these results were relevant to clinical problems involving excessive fear responses to stimuli that had not been previously paired with aversive experiences (e.g., excessive fear of photos of anything resembling a spider) and offered an alternative account to cognitive interpretations based on hypothetical constructs such as schema, beliefs, and expectancies.

As behavior analysts continue to study the nature of avoidance, scientists in other fields are producing better ways to predict what sorts of challenges—environmental as well as social—we will face in the future. Solutions to problems such as disease, poverty, and pollution depend on being able not only to predict their occurrence but also to change human behavior so as to avoid, to the extent possible, their disastrous outcomes. Thus, behavior analysts' contribution to an understanding of the fundamental determinants of avoidance and of ways to promote effective avoidance repertoires can directly influence our ability to prosper in an increasingly challenging world.

CHARACTERISTICS OF NEGATIVE REINFORCEMENT

Responses Acquired and Maintained by Negative Reinforcement

It is a well-known fact that aversive stimulation produces a variety of responses (Hutchinson, 1977). Some of these may be respondent behaviors (as in reflexive actions to intense stimuli), but the focus in this chapter is on operant behaviors. Recall that the presentation of an aversive stimulus serves as an EO for escape and occasions behavior that has produced escape from similar stimulation in the past. Any response that successfully terminates the stimulation will be strengthened; as a result, a wide range of behaviors may be acquired and maintained by negative reinforcement. All of these behaviors are adaptive because they enable one to interact effectively with the environment; some behaviors, however, are more socially appropriate than others. As will be seen later in the chapter, negative reinforcement may play an important role in the development of academic skills, but it also can account for the development of disruptive or dangerous behavior.

Events That Serve as Negative Reinforcers

In discussing the types of stimuli that can strengthen behavior through negative reinforcement, a problem arises when attempting to use the same terminology as that applied to the description of positive reinforcers. It is quite common to refer to positive reinforcers by listing stimuli such as food, money, praise, and so on. It is, however, the presentation of the stimulus that strengthens behavior: Food presentation, and not food per se, is a positive reinforcer. Nevertheless, we often simply list the stimulus and assume that "presentation" is understood. In a similar way, to say that negative reinforcers include shock, noise, parental nagging, and so on, is an incomplete description. It is important to remember that a stimulus described as a negative reinforcer refers to its removal because, as noted previously, the same stimulus serves as an EO when presented prior to behavior and as punishment when presented following behavior.

Learning History

As is the case with positive reinforcers, negative reinforcers influence behavior because (a) we have the inherited capacity to respond to them or (b) their effects have been established through a history of learning. Stimuli whose removal strengthens behavior in the absence of prior learning are **unconditioned negative reinforcers**. These stimuli typically are noxious events such as shock, loud noise, intense light, extremely high or low temperature, or strong pressure against the body. In fact, any source of pain or discomfort (e.g., a headache) will occasion behavior, and any response that successfully eliminates the discomfort will be reinforced. Other stimuli are **conditioned negative reinforcers**, which are previously neutral events that acquire their effects through pairing with an existing (unconditioned or conditioned) negative reinforcer. A bicyclist, for example, usually heads for home when seeing a heavily overcast sky because dark clouds have been highly correlated with impending bad weather. Various forms of social coercion, such as parental nagging, are perhaps the most commonly encountered conditioned negative reinforcers. For example, reminding a child to clean his or her bedroom may have little effect on the child's behavior unless failure to respond is followed by another consequence, such as having to stay in the room until it is clean. To the extent that the nagging is reliably "backed up" by sending the child to the room, the child will eventually respond simply to stop or prevent the nagging. It is interesting that, in the case of negative reinforcement, neutral events (dark sky, nagging) function as both (a) discriminative stimuli because responding in their presence avoids another consequence and (b) conditioned aversive stimuli (i.e., conditioned negative reinforcers) because, due to their pairing with another consequence, they become stimuli to avoid or escape.

The Source of Negative Reinforcement

Another way to classify negative reinforcers is based on how they are removed (i.e., the source of the stimulus change). In Chapter 11, a distinction was made between socially mediated reinforcement, in which the consequence results from the action of another person, and automatic reinforcement, in which the consequence is produced directly by a response independent of the actions of another. This distinction also applies to negative reinforcement. Returning to the example in Figure 12.1, we can see that termination of the construction noise was an instance of *social negative reinforcement* (the roommate's action closed the window). The person "being bothered" by the noise, however, could simply have walked across the room and closed the window (*automatic negative reinforcement*). This example illustrates the fact that many reinforcers can be removed or terminated either way: You can consult a physician when experiencing a headache (social) or take a pain medication (automatic), ask your roommate to turn off the stereo when you're studying for a test (social) or turn it off yourself (automatic), and so on.

Considering the source of negative reinforcement may facilitate the design of behavior-change procedures by determining the focus of intervention. For example, when faced with a perplexing work task, an employee may finish it incorrectly just to get it out of the way (automatic reinforcement) or ask for help (social reinforcement). Aside from reassigning the employee, the quickest solution would be to reinforce the employee's seeking help by offering assistance. Ultimately, however, the supervisor would want to teach the employee the necessary skills to complete the work tasks independently.

Identifying the Context of Negative Reinforcement

Chapter 11 outlined several ways to identify positive reinforcers; the difference with negative reinforcers is that equal emphasis must be placed on the antecedent event (EO) as well as on the reinforcing consequence because, once the behavior occurs, the negative reinforcer may be gone and cannot be observed. The identification of EOs may be difficult with individuals who have limited verbal ability and cannot tell someone they are experiencing aversive stimulation. These people may engage in other behaviors, such as tantrums, attempts to leave the situation, destructive behavior, aggression, or even self-injury. Weeks and Gaylord-Ross (1981), for example, observed students with severe disabilities when no

task, an easy task, and a difficult task were presented. Little or no problem behavior occurred during the no-task condition, and problem behavior occurred somewhat more often in the difficult-task condition than in the easy-task condition. These results suggested that the students' problem behavior was maintained by escape from task demands and that difficult tasks were more "aversive" than were easy tasks. However, because the consequences that followed problem behavior were unknown, it is possible that the behaviors were maintained by some other consequence, such as attention, which would be a positive reinforcer.

Iwata, Dorsey, Slifer, Bauman, and Richman (1994) developed a method for identifying the types of contingencies that maintain problem behavior by observing people under a series of conditions that differed with respect to both antecedent and consequent events. One condition involved the presentation of task demands (EO) and removal of the demands (escape) if problem behavior occurred; higher rates of problem behavior under this condition relative to others indicated that problem behavior was maintained by negative reinforcement (see Chapter 24 for further discussion of this approach to assessment).

Smith, Iwata, Goh, and Shore (1995) extended the findings of Weeks and Gaylord-Ross (1981) and Iwata et al. (1994) by identifying some characteristics of task demands that make them aversive. After first determining that their participants' (people with severe disabilities) problem behavior was maintained by escape from task demands, Smith and colleagues examined several dimensions along which tasks might differ: task novelty, duration of the work session, and rate of demand presentation. Results of one of these analyses are shown in Figure 12.4, which depicts frequency distributions and cumulative records of problem behavior from the beginning to the end of sessions. These data illustrate the importance of individualized assessments in

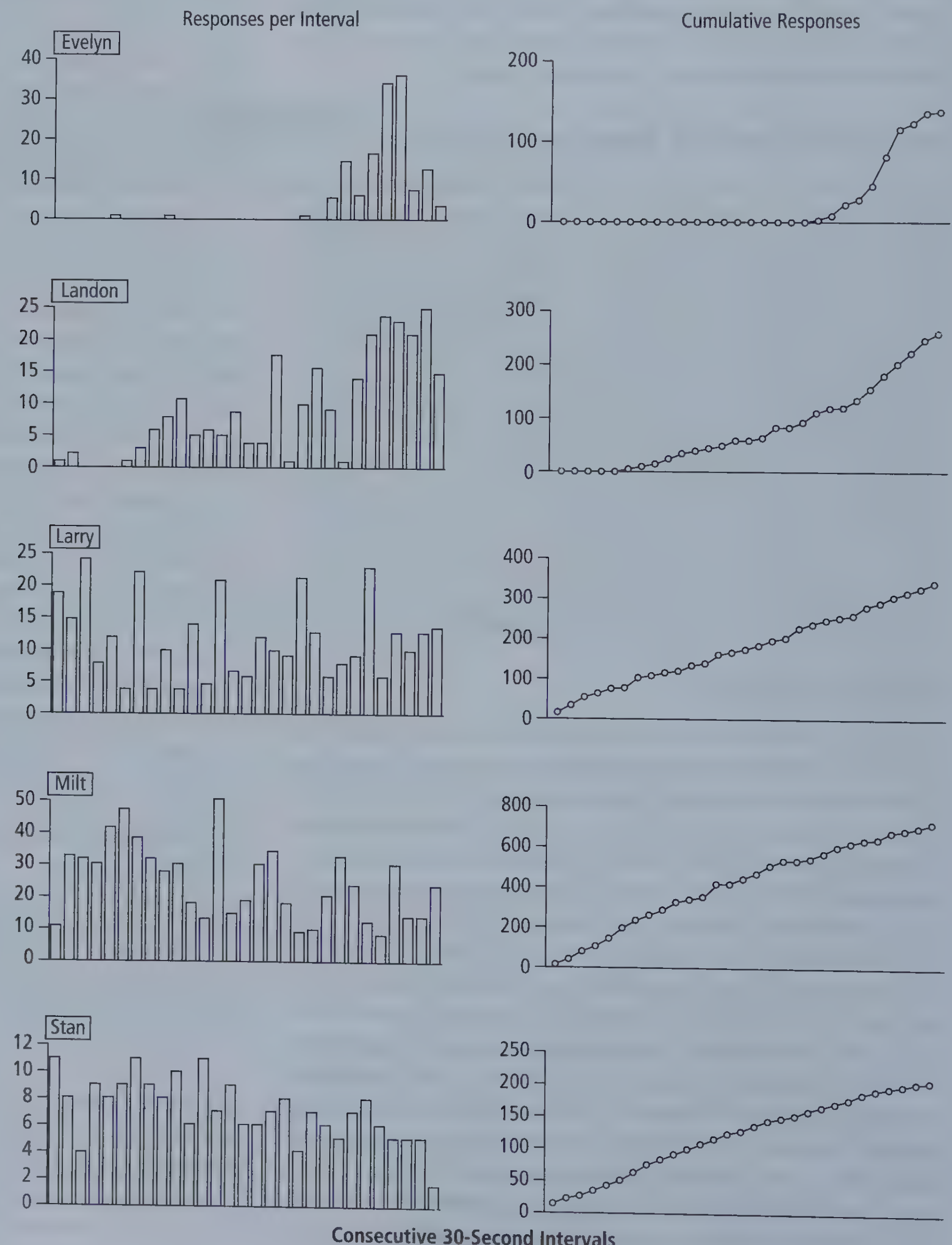


Figure 12.4 Frequency distributions (left column) and cumulative records (right column) summed over sessions of self-injurious behavior (SIB) by five adults with developmental disabilities as work sessions progressed.

identifying the basis for negative reinforcement because two participants (Evelyn and Landon) showed increasing rates of problem behavior as work sessions progressed, whereas two other participants (Milt and Stan) showed decreasing rates of problem behavior as work sessions progressed, as illustrated by the declining height of the bars in their histograms and of the slopes of their cumulative curves. In fact, subsequent studies (e.g., Roscoe, Rooker, Pence, & Longworth, 2009) have shown that failure to consider the idiosyncratic nature of aversive stimuli may lead to erroneous conclusions during assessment of problem behavior.

Determinants of Negative Reinforcement Effects

The factors that determine whether a negative reinforcement contingency will be effective in changing behavior are similar to those that influence positive reinforcement (see Chapter 11) and are related to (a) the strength of the contingency and (b) the presence of competing contingencies. In general, negative reinforcement for a given response will be more effective under the following conditions:

1. The stimulus change *immediately* follows the occurrence of the target response.
2. The *magnitude* of reinforcement is large, referring to the difference in stimulation present before and after the response occurs.
3. Occurrence of the target response *consistently* produces escape from or postponement of the EO.
4. Reinforcement is *unavailable* for competing (nontarget) responses.

APPLICATIONS OF NEGATIVE REINFORCEMENT

Negative reinforcement is a fundamental principle of learning that has been studied extensively in basic research (Hineline & Rosales-Ruiz, 2013). Although many examples of escape and avoidance learning can be found in everyday life, research in applied behavior analysis has heavily emphasized the use of positive reinforcement over negative reinforcement, mostly for ethical reasons, which are noted in the final section of this chapter. Still, negative reinforcement has been used as one means of establishing a variety of behaviors. This section illustrates several therapeutic uses of negative reinforcement, as well as the unintended role it may play in strengthening problem behavior.

Acquisition and Maintenance of Appropriate Behavior

Interventions for Chronic Food Refusal

Pediatric feeding problems are common and are especially prevalent among children with developmental disabilities. These disorders may take a variety of forms, including selective eating, failure to consume solid foods, and complete food refusal, and may be serious enough to require tube feeding or other artificial means to ensure adequate nutritional intake. A large proportion of feeding problems cannot be attributed to a medical cause but, instead, appear to be learned responses most likely maintained by escape or avoidance.

Results from a number of studies have shown that operant learning–based interventions can be highly effective in treating many childhood feeding disorders, and a study by Ahearn, Kerwin, Eicher, Shantz, and Swearingin (1996) illustrated the use of negative reinforcement as a form of intervention. Three children admitted to a hospital who had histories of chronic food refusal were first observed under a baseline (positive reinforcement) condition in which food was presented and access to toys was available contingent on accepting food. Food refusal, however, produced escape in that it terminated a trial. Subsequently, the experimenters compared the effects of two interventions. One treatment condition (nonremoval of the spoon) involved presenting food and keeping the spoon positioned at the child's lower lip until the bite was accepted. The other treatment (physical guidance) involved presenting food and, if the child did not accept, opening the child's mouth so that the food could be delivered. Both treatments involved a negative reinforcement contingency because food acceptance terminated the trial by producing removal of the spoon or avoidance of the physical guidance.

Figure 12.5 shows the results obtained for the three children. All children exhibited low rates of acceptance during baseline in spite of the availability of positive reinforcement. The two interventions were implemented in a multiple baseline across subjects design and were compared in a multielement design. As can be seen in the second phase of the study, both interventions produced immediate and large increases in food acceptance. Numerous subsequent studies have replicated these findings. Of particular interest is a study by Piazza, Patel, Gulotta, Sevin, and Layer (2003), who compared the effects of positive reinforcement (access to preferred toys and attention), negative reinforcement (physical guidance or nonremoval of the spoon, both labeled as "escape extinction" because food refusal no longer terminated feeding trials), and positive and negative reinforcement combined on chronic food refusal exhibited by four children. Results showed that positive reinforcement was not effective in increasing food consumption for any of the children. By contrast, negative reinforcement produced increases in consumption regardless of whether it was implemented alone or in combination with positive reinforcement. Collectively, results of these studies showed that positive reinforcement for appropriate behavior may have limited effects if competing behavior (food refusal) produces negative reinforcement, and that negative reinforcement, which maintained the children's problem behavior, could be used to establish alternative behavior.

Error-Correction Strategies

As noted in Chapter 11, positive reinforcement is a basic motivational component of effective instruction. Teachers commonly deliver praise, privileges, and other forms of reward contingent on correct performance. Another common procedure, but one that has received less attention than positive reinforcement, involves the correction of student errors by repeating a learning trial, having the student practice correct performance, or giving the student additional work. To the extent that correct performance avoids these remedial procedures, improvements may be just as much a function of negative reinforcement as positive reinforcement.

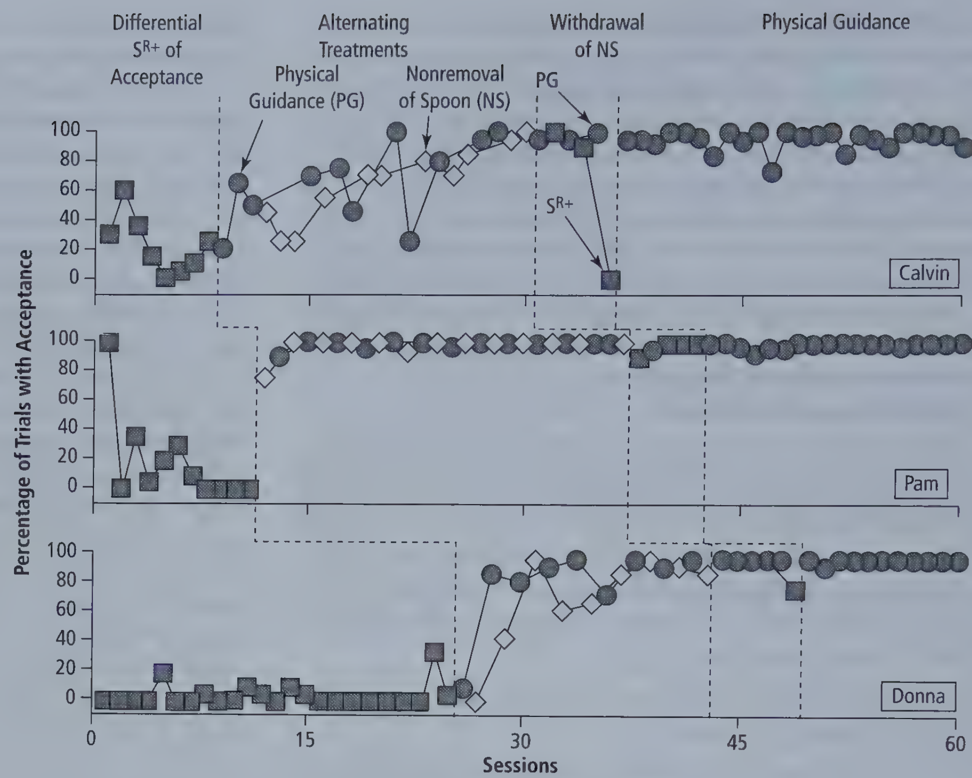


Figure 12.5 Percentage of trials in which three children with histories of chronic food refusal accepted bites during a baseline condition of positive reinforcement and two treatment conditions, nonremoval of the spoon and physical guidance, both of which involved a negative reinforcement contingency.

From "An Alternating Treatments Comparison of Two Intensive Interventions for Food Refusal" by W. H. Ahearn, M .E. Kerwin, P. S. Eicher, J. Shantz, and W. Swearingin, 1996, *Journal of Applied Behavior Analysis*. Reproduced with permission of John Wiley &Sons, Inc.

Worsdell and colleagues (2005) examined the relative contributions of these contingencies during behavioral acquisition. The learning task involved reading words presented on flash cards, and the intervention of interest involved the correct repetition of misread words during remedial trials. The procedure provided additional practice of correct responses but also represented an avoidance contingency. To separate these effects (in Study 3), the authors implemented two error correction conditions. In the "relevant" condition, which combined the effects of practice and negative reinforcement, students were prompted to pronounce the misread word correctly 5 times contingent on an error. In the "irrelevant" condition, students were prompted to repeat an unrelated, nontarget word 5 times contingent on an error. The irrelevant condition contained only the negative reinforcement contingency because repetition of irrelevant words provided no practice in correctly reading misread words.

Figure 12.6 shows the results of Study 3, expressed as the cumulative number of words mastered by the 9 participants. All participants' performance improved during both error-correction conditions relative to baseline, when no error-correction procedure was in effect. Performance by 3 participants (Tess, Ariel, and Ernie) was better during relevant error correction. However, Mark's performance was clearly superior during irrelevant error correction, and performance of the remaining 5 participants (Hayley, Becky, Kara, Maisy, and Seth) was similar in both conditions. Thus, all participants showed improvement in reading performance even when they practiced irrelevant words, and most participants (6 of 9) did just as well or better practicing irrelevant rather than relevant words. These results suggest that the success of many remedial (error-correction) procedures may be due at least in part to negative reinforcement. Subsequent studies (Carroll, Joachim, St. Peter, & Robinson, 2015; Kodak et al., 2016; McGhan & Lerman, 2013) have examined specific features of error-correction procedures that may enhance their effectiveness, and a consistent finding has been the contribution of avoidance contingencies to learning in instructional contexts. That is, the inclusion of a remedial-trial contingency produces faster learning than that resulting from positive reinforcement of correct responses alone.

Acquisition and Maintenance of Problem Behavior

Well-designed instructional procedures maintain a high degree of task engagement and lead to improved learning. Occasionally, however, the presentation of task demands may function as an EO for escape behavior due to some aversive feature of the work requirements. Initial forms of escape may include lack of engagement or mild forms of disruption. To the extent that positive reinforcement for compliance is less than optimal, attempts to escape may persist and may escalate to severe forms of problem behavior. In fact, research on the assessment and treatment of problem behavior has shown that escape from task demands is a common source of negative reinforcement for property destruction, aggression, and even self-injury. This topic is covered more extensively in Chapter 27 and is included here as well because of its special relevance to negative reinforcement.

O'Reilly (1995) conducted an assessment of a person's episodic aggressive behavior. The participant was an adult with severe intellectual disability who attended a vocational day program. To determine whether aggressive behavior was maintained by positive versus negative reinforcement, O'Reilly observed the participant under two conditions, which were alternated in a multielement design. In one condition (attention), a

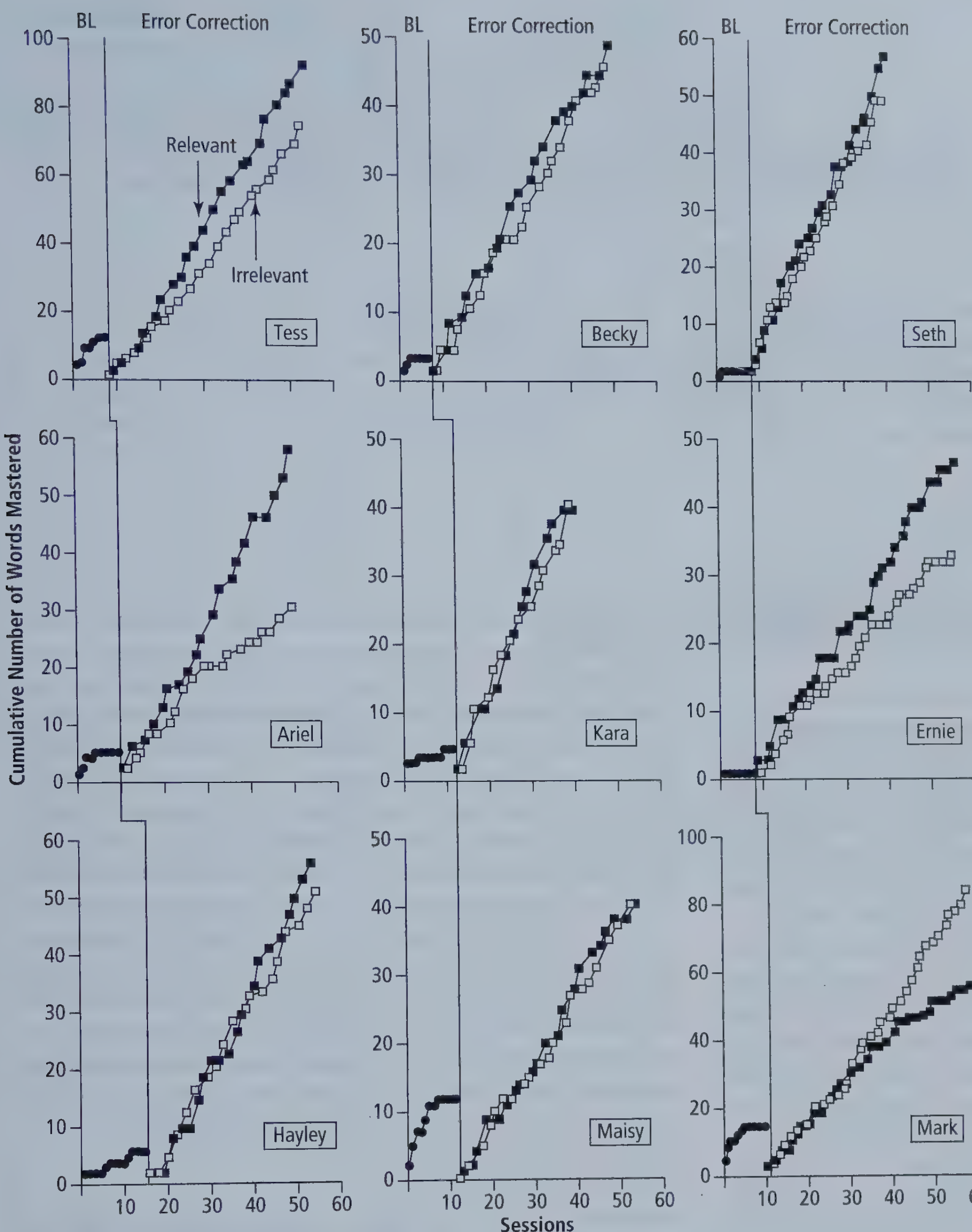


Figure 12.6 Cumulative number of words read correctly during baseline (positive reinforcement for correct responses) and during two error-correction conditions: one in which correct responses avoided repeated practice of the misread (relevant) word, and another in which correct responses avoided practice of an unrelated (irrelevant) word. Improvements in performance in the "irrelevant" condition indicated that negative reinforcement plays a role in error-correction procedures.

"Analysis of Response Repetition as an Error-Correction Strategy During Sight-Word Reading" by A. S. Worsdell, B. A. Iwata, C. L. Dozier, A. D. Johnson, P. L. Neidert, and J. L. Thomason, 2005, *Journal of Applied Behavior Analysis*, *38*, p. 524. Copyright 2005 by the Society for the Experimental Analysis of Behavior, Inc.

therapist ignored the participant (EO) except to deliver reprimands following aggression (positive reinforcement). In the second condition (demand), a therapist presented difficult tasks to the participant (EO) and briefly terminated the trial following aggression (negative reinforcement).

As Figure 12.7 shows, aggressive behavior occurred more often in the demand condition, indicating that it was maintained by negative reinforcement. Because anecdotal reports suggested that the participant also was more likely to be aggressive following nights when he had not slept well, the data for both conditions were further divided based on whether the participant slept for more or less than 5 hours the previous night. The highest rates of aggression occurred following sleep deprivation. These data are particularly interesting in that they illustrate the combined influence of two antecedent events—EOs on behavior maintained by negative reinforcement: Work tasks (EO #1) generally occasioned escape but even more so in the absence of sleep (EO #2).

Other studies have had similar outcomes, suggesting that a number of "contextual" factors can mediate the "aversiveness" of demand situations. For example, task demands may be more likely to evoke escape behavior when a medical condition such as otitis media (ear infection) is present (O'Reilly, 1997) or when demands are presented in the context of low-quality attention (e.g., negative verbal statements) from caregivers (Gardner, Wacker, & Boelter, 2009). By contrast, contextual manipulations such as presenting the opportunity to choose among tasks (Dyer, Dunlap, & Winterling, 1990) or to choose the order in which tasks are completed (Kern, Mantegna, Vorndran, Bailin, & Hilt, 2001), or presenting demands in the context of pleasant stories (Carr, Newsom, & Binkoff, 1976), have been shown to decrease problem behaviors maintained by escape.

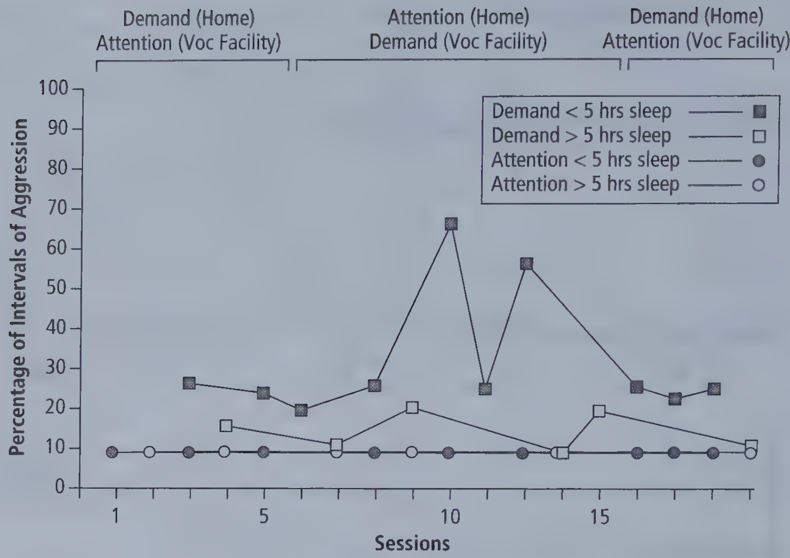


Figure 12.7 Data showing that the effects of work tasks as EOs for escape-maintained aggression by an adult male with severe intellectual disability were exacerbated by sleep deprivation.

"Functional Analysis and Treatment of Escape-Maintained Aggression Correlated with Sleep Deprivation" by M. F. O'Reilly, 1995, *Journal of Applied Behavior Analysis, 28*, p. 226. Copyright 1995 by the Society for the Experimental Analysis of Behavior, Inc.

Extinction of Behavior Maintained by Negative Reinforcement

When operant behavior no longer produces reinforcement, its frequency decreases through a process known as extinction. As is the case with reinforcement, extinction involves both an operation (termination of reinforcement) and an effect (a reduction in responding). Although simple in concept, the application of extinction requires consideration of numerous factors: (a) its functional as well as procedural variations; (b) its direct and indirect effects, some of which may be undesirable; and (c) historical influences that may facilitate or impede its outcome. As a result, a chapter of this text (Chapter 24) is devoted to the topic of extinction. We introduce it here just to note that extinction represents one way (the most direct way) to decrease behavior maintained by negative reinforcement and to emphasize the fact that termination of a negative reinforcement contingency by "no longer reinforcing behavior" means presenting the consequence that responding previously escaped or avoided (i.e., withholding negative reinforcement). For example, if a child's loud whining and complaining allow the child to delay preforming an undesirable task or perhaps to avoid it altogether, extinction would involve reinstating the requirement to complete the task.

Behavioral Replacement Strategies

Differential Negative Reinforcement

Problem behaviors maintained by negative reinforcement can be treated in a number of ways. One strategy is to strengthen a more socially appropriate replacement behavior using negative reinforcement, as illustrated in a study by Durand and Carr (1987). After determining that the "stereotypic" behaviors of four special education students were maintained by escape from task demands, the authors taught the students an alternative response ("Help me"), which was followed by assistance with the task at hand. As can be seen in Figure 12.8, all students engaged in moderate-to-high levels of stereotypy during baseline. After being taught to use the phrase "Help me," the students began to exhibit that behavior, and their stereotypy decreased.

Results of the Durand and Carr (1987) study showed that a desirable behavior could replace an undesirable one; however, the replacement behavior might be considered less than ideal because it did not necessarily facilitate improved task engagement. This limitation was addressed in a subsequent study by Marcus and Vollmer (1995). After collecting baseline data on a young girl's compliant and disruptive behavior, the authors compared the effects of two treatments in a reversal design. In a differential negative reinforcement (DNR) communication condition, the girl was given a brief break from the task when she said "Finished." In a second condition, called DNR compliance, the girl was given a break after complying with an instruction (the criterion for a break was later increased to compliance with three instructions). Results of this comparison (see Figure 12.9) showed that both treatments produced marked reductions in disruptive behavior. However, only the DNR compliance condition produced an improvement in task performance.

Differential Positive Reinforcement

Research on the treatment of problem behavior has emphasized the importance of extinction as a critical component of intervention. Results of a small but growing number of studies, however, indicate that reinforcement-based procedures occasionally can be effective even though problem behavior might continue to be reinforced. Furthermore, some of these studies have shown that applications of positive reinforcement can be successful in strengthening appropriate replacement behavior when problem behavior has been maintained by negative reinforcement (see Payne & Dozier, 2013, for a recent review).

A study by Slocum and Vollmer (2015) illustrated this approach with five children who engaged in high rates of problem behavior maintained by escape from task demands and correspondingly low levels of compliance. The authors compared two interventions in a multielement design, both of which were implemented without extinction (i.e., problem behavior continued to produce escape). Positive reinforcement involved delivery of a small edible item contingent on compliance, whereas negative reinforcement involved delivering a 30-sec break from the task contingent on compliance. The authors reported

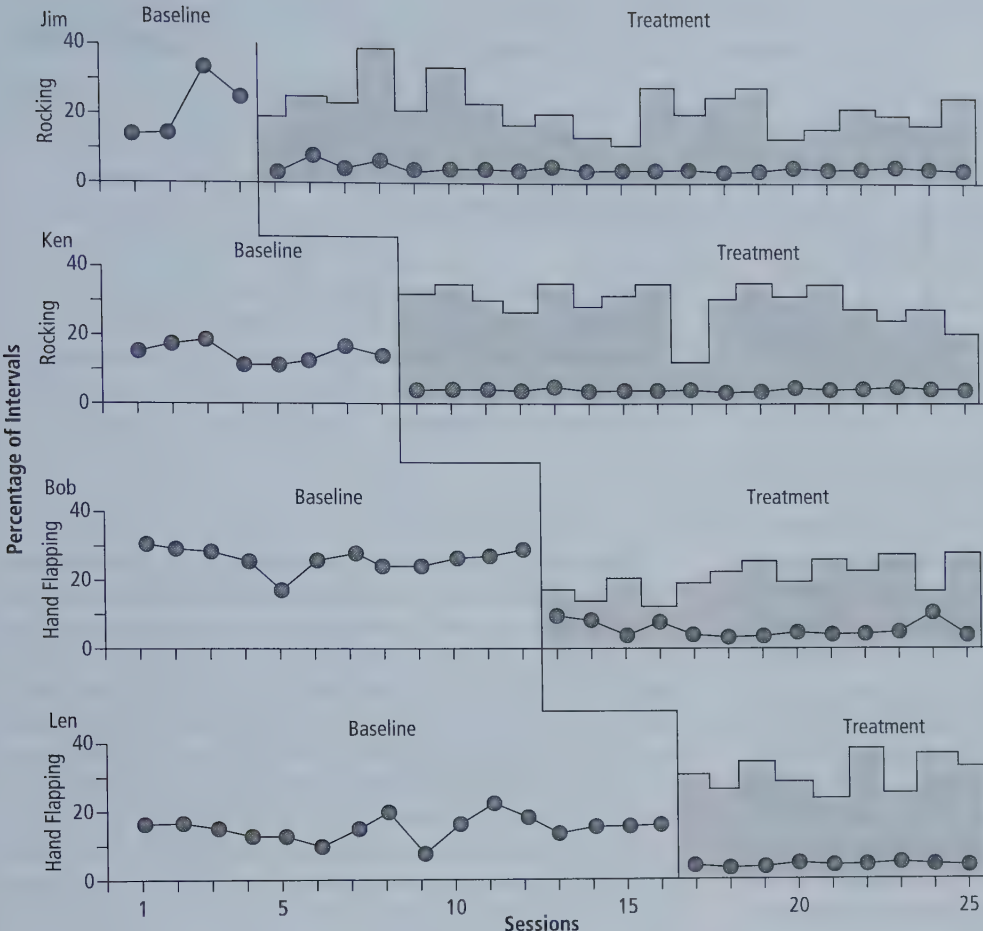


Figure 12.8 Percentage of intervals of stereotypic behaviors maintained by escape from task demands of four special education students during baseline and treatment in which the students were taught an alternative response ("Help me") to attain assistance with the task at hand. Shaded bars show students' use of the "Help me" response.

"Social Influences on 'Self-Stimulatory' Behavior: Analysis and Treatment Application" by V. M. Durand and E. G. Carr, E. G., 1987, *Journal of Applied Behavior Analysis*, *20*, 128. Copyright 1987 by the Society for the Experimental Analysis of Behavior, Inc.

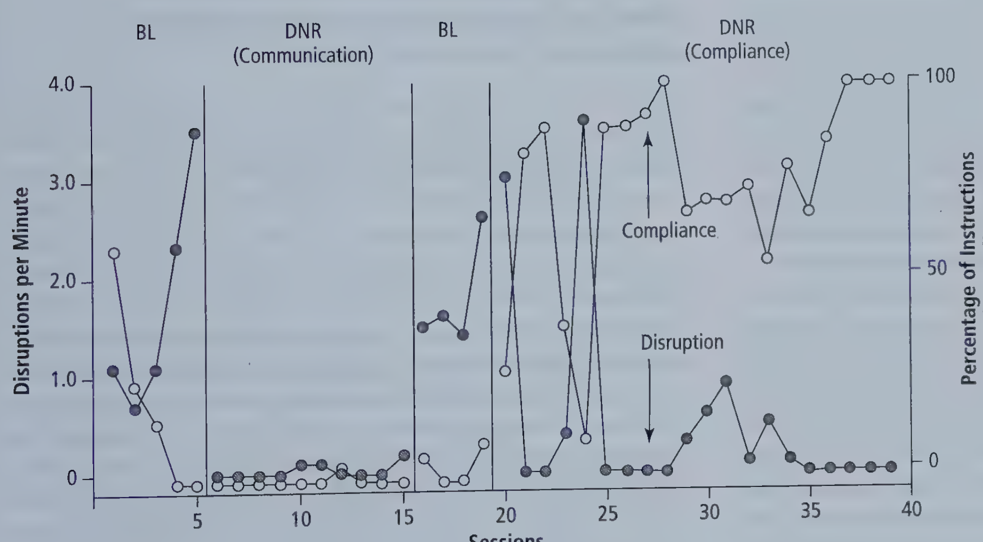


Figure 12.9 Disruptions and compliance by a 5-year-old girl during baseline and two differential negative reinforcement conditions.

"Effects of Differential Negative Reinforcement on Disruption and Compliance" by B. A. Marcus and T. R. Vollmer, 1995, *Journal of Applied Behavior Analysis*, *28*, p. 230. Copyright 1995 by the Society for the Experimental Analysis of Behavior, Inc.

that positive reinforcement produced more consistent effects, as illustrated in Figure 12.10, which shows data on problem behavior for one of the participants. During the first treatment phase, both interventions were effective, although lower rates of problem behavior were observed under the positive reinforcement condition. During the second treatment phase, positive reinforcement was effective, whereas negative reinforcement was not.

CHANGES IN TEACHER AND CAREGIVER RESPONDING AS A FUNCTION OF NEGATIVE REINFORCEMENT

A great deal of research has focused on the use of procedures such as those described above to improve the behavior of students and clients. By contrast, very little work has examined the complementary effect of client behavior change on the behavior

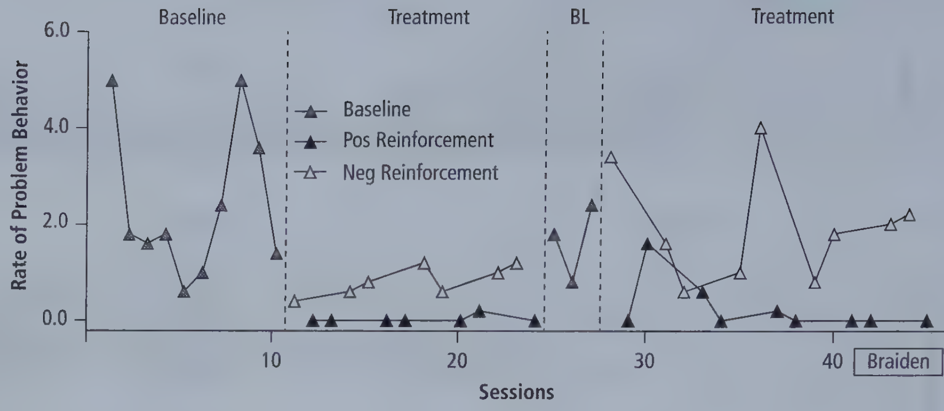


Figure 12.10 The rate of problem behavior during baseline and treatment conditions (left) and the percentage of compliance during baseline and treatment conditions (right) across sessions. (*Note*: This graph shows results for one of the study' five subjects).

From "A comparison of positive and negative reinforcement for compliance to treat problem behavior maintained by escape." by S. K. Slocum and T. R. Vollmer, 2015, *Journal of Applied Behavior Analysis*. Reproduced with permission of John Wiley & Sons,Inc.

of the agent (teacher, therapist, etc.), yet this relation may reflect an important negative reinforcement contingency. When a teacher implements a behavior change procedure that results in a reduction of problem behavior, teacher behavior has terminated an aversive event. This type of relation may strengthen any behavior—inappropriate as well as appropriate—that has the same effect, a possibility suggested by results from descriptive studies of teacher behavior (Addison & Lerman, 2009; Sloman et al., 2005).

Negative reinforcement effects have been examined more directly in two experimental simulations. Miller, Lerman, and Fritz (2010) observed seven teachers in training who were instructed to teach a variety of skills to a "student" (role played by a research assistant). The student engaged in problem behavior at prearranged times throughout sessions and either stopped when the teacher delivered a reprimand (negative reinforcement condition) or continued to engage in problem behavior regardless of the teacher's behavior (extinction). Results showed that student behavior had a clear effect on teacher behavior in that the rate of reprimands was higher during the negative reinforcement condition. The authors also reported that three of the teachers terminated their participation in the study during an extinction condition, which provided further evidence that "problem behavior" was an aversive event. Using a similar methodology, Thompson, Bruzek, and Cotnoir-Bichelman (2011) observed 11 undergraduate students during simulated caretaking sessions in which the student interacted with a "baby" (a doll). An experimenter presented a recorded infant cry at the beginning of sessions and terminated the cry contingent on different caregiver responses (horizontal rocking, vertical rocking, simulated feeding, playing) in different conditions. Results generally showed persistence of responses that were "effective" in terminating crying and extinction of responses that no longer were effective.

Data from this line of research have several implications. First, the occurrence of problem behavior appears to be an EO for caregiver escape and avoidance and is likely to occasion any caregiver response that is effective in terminating it, ranging from active (delivering a reprimand) to passive (limiting social interaction) responding. Second, this fact underscores the importance of instructing parents and teachers in appropriate behavior management strategies to prevent the emergence of inappropriate ones. Finally, the maintenance of newly taught therapeutic skills may be more than a function of the quality of training; an equally important determinant may be how successful those skills are in quickly solving problems.

ETHICAL CONSIDERATIONS IN USING NEGATIVE REINFORCEMENT

Ethical concerns about the use of positive and negative reinforcement arise primarily from the severity of the antecedent event (EO) that occasions the target behavior. Many EOs for behavior maintained by positive reinforcement can be characterized as deprivation states, which, if severe, can constitute undue restriction of rights. By contrast, EOs for behavior maintained by negative reinforcement are, by definition, aversive events. Extremely noxious events, when presented as antecedent stimuli, cannot be justified as part of a typical behavior change program.

Another concern with negative reinforcement is that the presence of aversive stimuli can itself generate behaviors that compete with the acquisition of desired behavior (Hutchinson, 1977; Myer, 1971). For example, a socially withdrawn child, when placed in the midst of others, may simply scream and run away instead of playing with the peers, and running away is incompatible with social interaction. Finally, undesirable side effects typically associated with punishment (see Chapter 14) might also be observed when implementing behavior change programs based on negative reinforcement.

SUMMARY

Definition of Negative Reinforcement

1. Negative reinforcement involves the termination, reduction, or postponement of a stimulus contingent on the occurrence of a response, which leads to an increase in the future occurrence of that response.

2. A negative reinforcement contingency involves (a) an establishing operation (EO) in whose presence escape is reinforcing, (b) a discriminative stimulus (S^D) in whose presence a response is more likely to be reinforced, (c) the response that produces reinforcement, and (d) termination of the event that served as the EO.

3. Positive and negative reinforcement are similar in that both lead to an increase in responding; they differ in that positive reinforcement involves contingent stimulus presentation, whereas negative reinforcement involves contingent stimulus termination.

4. Negative reinforcement and punishment differ in that (a) negative reinforcement involves contingent stimulus termination, whereas punishment involves contingent stimulation, and (b) negative reinforcement leads to an increase in responding, whereas punishment leads to a decrease in responding.

Escape and Avoidance Contingencies

5. An escape contingency is one in which responding terminates an ongoing stimulus. An avoidance contingency is one in which responding delays or prevents the presentation of a stimulus.

6. In discriminated avoidance, responding in the presence of a signal prevents stimulus presentation; in free-operant avoidance, responding at any time prevents stimulus presentation.

Characteristics of Negative Reinforcement

7. Any response that successfully terminates aversive stimulation will be strengthened; as a result, a wide range of behaviors may be acquired and maintained by negative reinforcement.

8. Negative reinforcement may play an important role in the development of academic skills, but it can also account for the development of disruptive or dangerous behavior.

9. Unconditioned negative reinforcers are stimuli whose removal strengthens behavior in the absence of prior learning. Conditioned negative reinforcers are stimuli whose removal strengthens behavior as a result of previous pairing with other negative reinforcers.

10. Social negative reinforcement involves stimulus termination through the action of another person. Automatic negative reinforcement involves stimulus termination as a direct result of a response.

11. Identification of negative reinforcers requires the specification of the stimulus conditions in effect prior to and following responding.

12. In general, negative reinforcement for a given response will be more effective when (a) the stimulus change immediately follows the occurrence of the target response, (b) the magnitude of reinforcement is large, (c) the target response consistently produces escape from or postponement of the EO, and (d) reinforcement is unavailable for competing responses.

Applications of Negative Reinforcement

13. Although negative reinforcement is a fundamental principle of learning that has been studied extensively in basic research, applied behavior analysis has heavily emphasized the use of positive reinforcement over negative reinforcement.

14. Applied researchers have explored the therapeutic uses of negative reinforcement in treating pediatric feeding problems.

15. Improvements in student performance as a result of error correction that involves repeating a learning trial, having the student practice correct performance, or giving the student additional work may be a function of negative reinforcement.

16. The presentation of task demands during instruction may function as an EO for escape; initial forms of escape may include lack of attention or mild forms of disruption. To the extent that positive reinforcement for compliance is less than optimal, escape behaviors may persist and may even escalate.

17. Extinction of behavior maintained by negative reinforcement involves *not* allowing a response to avoid or terminate its programmed consequence.

18. One strategy for treating problem behaviors maintained by negative reinforcement is to strengthen a more socially appropriate replacement behavior through negative reinforcement.

19. Results of some research have shown that interventions based on positive reinforcement also can be effective as treatment for problem behavior maintained by negative reinforcement.

Changes in Teacher and Caregiver Responding as a Function of Negative Reinforcement

20. Another dimension of negative reinforcement involves effects on caregiver behavior. Student problem behavior serves as an EO teacher escape and avoidance, and teacher responses to problem behavior are maintained by negative reinforcement.

Ethical Considerations in Using Negative Reinforcement

21. Ethical concerns about the use of positive and negative reinforcement are similar and arise from the severity of the antecedent event (EO) that occasions behavior. Most EOs for behavior maintained by negative reinforcement can be viewed as aversive events. Extremely noxious events, when presented as antecedent stimuli, cannot be justified as part of a typical behavior change program.

22. Another concern with negative reinforcement is that the presence of aversive stimuli can itself generate behaviors that compete with the acquisition of desired behavior.

KEY TERMS

avoidance contingency
conditioned negative reinforcer
discriminated avoidance
escape contingency
free-operant avoidance
negative reinforcement
unconditioned negative reinforcer

MULTIPLE-CHOICE QUESTIONS

1. Negative reinforcement can be defined as: Stimulus _______________, contingent upon a response, which _______________ the future probability of that response.

a. Presented, increases

b. Presented, decreases

c. Removed, increases

d. Removed, decreases

Hint: (See "Definition of Negative Reinforcement")

2. Which of the following is an example of a negative reinforcement contingency?

a. Jo is sitting near a window at Starbucks and having a cup of coffee. The sun is streaming in the window, and it is too warm for Jo—she is beginning to perspire. Jo moves to another chair away from the window, where it is shady. The next time Jo goes to Starbucks, she sees the sun shining in the window again and sits in the chair in the shade instead.

b. Frank often goes to Starbucks for coffee as well. After receiving his coffee, he sits down in his seat. It is 9:00 am, and Starbucks is very busy. Most of the tables are full. As a result, a smart woman approaches Frank and asks if she can share his table. Frank agrees, and the woman sits down and has coffee with Frank. Frank decides he will come to Starbucks at 9:00 am for coffee more often in the future.

c. Cheryl loves coffee. She orders a coffee at Starbucks but burns her mouth on it because it is too hot. After a couple of occurrences of this, Cheryl stops coming to Starbucks for coffee.

d. Sam was walking to Starbucks for coffee. On his way, he lost his wallet to a pickpocket. Sam no longer walks to that Starbucks for coffee.

Hint: (See "Definition of Negative Reinforcement")

3. Mary Jo decides to implement a negative reinforcement intervention with one of her students, Anjali, to increase the amount of work Anjali completes. Mary Jo tells Anjali, "If you complete 15 math problems today (Monday), you don't have to do your math worksheet on Friday." What is one problem that might arise with this intervention?

a. Reinforcement is available for a competing behavior.

b. There is little difference in the stimulus conditions before the response (completing work) occurs and after the response occurs.

c. The stimulus change following the occurrence of target behavior is not immediate.

d. Mary Jo is being inconsistent in her delivery of reinforcement.

Hint: (See "Determinants of Negative Reinforcement Effects")

4. Unlike assessments for identifying positive reinforcers, assessments for negative reinforcers must place equal emphasis on _______________ as well as the consequence events for target behavior.

a. The EO or antecedent event

b. The future occurrence of the target behavior

c. The stimulus

d. Punishment

Hint: (See "Identifying the Context of Negative Reinforcement")

5. The key difference between an escape contingency and an avoidance contingency is:

a. In an escape contingency, the future probability of the target behavior increases, while the future probability of the target behavior decreases with an avoidance contingency.

b. In an escape contingency, the EO is present prior to the occurrence of the target behavior, while in an avoidance contingency, the EO is not present prior to the occurrence of the target behavior.

c. In an escape contingency, the EO is not present prior to the occurrence of the target behavior, while in an avoidance contingency, the EO is present prior to the occurrence of the target behavior.

d. In an escape contingency, there is a warning stimulus that signals an aversive stimulus is imminent, while in an avoidance contingency, there is no such warning stimulus.

Hint: (See "Escape and Avoidance Contingencies")

6. Which of the following is an example of free-operant avoidance?

a. Jackie crosses the street when she sees Donna come around the corner on the next block so that she doesn't have to talk to her.

b. Lavonda puts on a bicycle helmet when she rides her bike so that she doesn't hurt her head if she falls off her bike.

c. Nathaniel puts down the hurricane shutters on his Florida home when he hears that a tropical storm is approaching so that his windows do not break.

d. All of these

Hint: (See "Escape and Avoidance Contingencies")

7. Which of the following behaviors could be maintained by negative reinforcement?

a. Completing school work.

b. Cleaning a bedroom

c. Tantrums

d. All of these

Hint: (See "Characteristics of Negative Reinforcement")

8. An unconditioned negative reinforcer:

a. Is one that began as a neutral stimulus

b. Acquired its aversive qualities by being paired with another aversive stimulus

c. Is one that strengthens behavior in the absence of prior learning

d. Can be thought of as an inherited negative reinforcer

e. Third and fourth choices

Hint: (See "Characteristics of Negative Reinforcement")

9. The textbook describes a study by Ahearn and colleagues (1996), in which negative reinforcement was used to increase food acceptance in children. In this example, during baseline, bite acceptances produced access to toys and bite refusals produced removal of the spoon (negative reinforcement). During the intervention, bite refusals no longer produced removal of the spoon. Instead, the spoon was only removed if a bite was accepted. As soon as a bite was accepted and every time a bite was accepted, the spoon was briefly removed. Which factors that are important to consider for effectively changing behavior with negative reinforcement are illustrated in this example?

a. The stimulus change following the occurrence of the target behavior was immediate.

b. The difference in stimulation prior to and after the response occurred was large.

c. The occurrence of the target response consistently produced escape.

d. Reinforcement was unavailable for competing responses.

e. All of these

f. None of these

Hint: (See "Factors That Influence the Effectiveness of Negative Reinforcement")

10. The study by Rodgers and Iwata (1991) that analyzed the effects of positive reinforcement, error correction procedures, and an avoidance procedure demonstrated that:

a. Error correction procedures may produce learning, at least in part, due to an avoidance contingency.

b. Positive reinforcement was the most effective procedure for producing correct performance.

c. Error correction procedures and avoidance procedures produced the most correct responding.

d. All of these

e. None of these

Hint: (See "Applications of Negative Reinforcement")

11. Ethical concerns about the use of negative reinforcement stem from:

a. Having to deprive the individual of positive reinforcement for an extended period of time

b. The presence of antecedent aversive stimuli in the individual's environment

c. The potential for creating a context that generates undesired behaviors

d. The presentation of aversive stimuli contingent upon a target behavior displayed by an individual

e. Second and third choices

Hint: (See "Ethical Considerations in Using Negative Reinforcement")

12. The potential negative side effects (e.g., crying, running away) of negative reinforcement are similar to the side effects associated with:

a. Positive Reinforcement

b. Punishment

c. Stimulus Control

d. All of these

e. None of these

Hint: (See "Ethical Considerations in Using Negative Reinforcement")

ESSAY-TYPE QUESTIONS

1. Write the definition of negative reinforcement and provide an example of negative reinforcement that has not been discussed in class or presented in the textbook. When giving your example, write it in the form of the diagrams found in the textbook, and be sure to include all four components of the diagram.

   Hint: (See "Definition of Negative Reinforcement")

2. What are some stimuli that might serve as negative reinforcers? Identify at least one negative reinforcer that is unconditioned and one negative reinforcer that is conditioned. Then, select one and show how that negative reinforcer could be used to teach a behavior (the behavior can either be socially appropriate or socially inappropriate). Use the diagramming notation presented in class to illustrate the negative reinforcement contingency.

   Hint: (See "Characteristics of Negative Reinforcement")

3. What is the key difference between an escape contingency and an avoidance contingency?

   Hint: (See "Escape and Avoidance Contingencies")

4. Explain the difference between social negative reinforcement and automatic negative reinforcement.

   Hint: (See "Characteristics of Negative Reinforcement")

5. What are four things you can do to increase the effectiveness of negative reinforcement?

   Hint: (See "Determinants of Negative Reinforcement Effects")

6. You are working with a child with autism, who engages in problem behavior to escape completing daily living tasks, such as washing his hands before lunch. You would like to teach this student to politely say, "Don't want to wash hands," rather than to throw tantrums when told to wash his hands. You set up an instructional program where you will make several hand-washing requests throughout the day in order to provide the child with opportunities to practice saying "Don't want to wash hands." Each time the child makes this request, you allow him to escape washing his hands. What are the ethical issues that might arise from such a plan and how can you plan for and/or decrease these negative effects?

   Hint: (See "Ethical Issues in Using Negative Reinforcement")

NOTES

1. The term *aversive* is not meant to describe an inherent characteristic of a stimulus but, rather, a stimulus whose presentation functions as punishment or whose removal functions as negative reinforcement.

CHAPTER 13

Schedules of Reinforcement

LEARNING OBJECTIVES

- Define schedule of reinforcement.
- Define continuous reinforcement.
- Define intermittent reinforcement.
- Be able to identify the importance of naturally occurring reinforcement.
- Identify the two main intermittent schedules of reinforcement.
- Compare and contrast fixed and variable ratio schedules.
- Explain the fixed ratio schedule consistency of performance.
- Explain the phenomenon of a post-reinforcement pause.
- Explain and define a variable ratio.
- Explain and describe fixed interval schedules of reinforcement.
- Explain and define variable interval schedules of reinforcement.
- Be able to define and explain the variables associated with "ratio strain."
- Identify four variations of basic intermittent schedules of reinforcement.

A **schedule of reinforcement** is a rule that describes a contingency of reinforcement, the environmental arrangements that determine conditions by which behaviors will produce reinforcement. Continuous reinforcement and extinction provide the boundaries for all other schedules of reinforcement. A schedule of **continuous reinforcement (CRF)** provides reinforcement for each occurrence of behavior. For example, a teacher using a continuous schedule of reinforcement would praise each occurrence of a correctly identified sight word emitted by a student. Examples of behaviors that tend to produce continuous reinforcement include turning on a water faucet (water comes out), turning on a light switch in a dark room (a light comes on), and putting money into a vending machine (a product is obtained). During extinction (EXT), no occurrence of the behavior produces reinforcement. (For a detailed description of extinction, see Chapter 24.)

INTERMITTENT REINFORCEMENT

Between continuous reinforcement and extinction, many **intermittent schedules of reinforcement (INT)** are possible in which some, but not all, occurrences of the behavior are reinforced. Only selected occurrences of behavior produce reinforcement with an intermittent schedule of reinforcement. CRF is used to strengthen behavior, primarily during the initial stages of learning new behaviors. INT is used to maintain an established behavior, especially during maintenance stages of learning.

Maintenance of Behavior

Maintenance of behavior refers to a lasting change in behavior. Regardless of the type of behavior change technique employed or the degree of success during treatment, applied behavior analysts must be concerned with sustaining gains after terminating a treatment program. For example, Zofia is in the seventh grade and taking French, her first foreign language class. After a few weeks, the teacher informs Zofia's parents that she is failing the course. The teacher believes that Zofia's problems in French have resulted from lack of daily language practice and study. The parents and teacher decide that Zofia will record a tally on a chart kept on the family bulletin board each evening that she studies French for 30 minutes. Zofia's parents praise her practice and study accomplishments and offer encouragement. During a follow-up meeting 3 weeks later, the parents and teacher decide that Zofia has done so well that the tally procedure can be stopped. Unfortunately, a few days later Zofia is once again falling behind in French.

A successful program was developed to establish daily French language practice. However, gains were not maintained after removing the tally procedure. The parents and the teacher did not establish intermittent reinforcement procedures. Let us review what happened and what could have happened. Continuous reinforcement was used correctly to develop daily study behavior. However, after the study behavior was established and the tally procedure removed, the parents should have continued to praise and encourage daily practice and

gradually offer fewer encouragements. The parents could have praised Zofia's accomplishments after every second day of daily practice, then every fourth day, then once per week, and so on. With the intermittent praise, Zofia might have continued daily practice after removing the tally procedure.

Progression to Naturally Occurring Reinforcement

A major goal of most behavior change programs is the development of naturally occurring activities, stimuli, or events to function as reinforcement. It is more desirable for people to read because they like to read, rather than to obtain contrived reinforcement from a teacher or parent; to engage in athletics for the enjoyment of the activity, rather than for a grade or because of a physician's directive; to help around the house for the personal satisfaction it brings, rather than to earn an allowance. Intermittent reinforcement is usually necessary for the progression to naturally occurring reinforcement. Even though some individuals spend hours each day practicing a musical instrument because they enjoy the activity, chances are good that this persistent behavior developed gradually. At first, the beginning music student needs a great deal of reinforcement to continue the activity: "You really practiced well today," "I can't believe how well you played," "Your mother told me you received a first place in the contest—that's super!" These social consequences are paired with other consequences from teachers, family members, and peers. As the student develops more proficiency in music, the outside consequences occur less frequently, intermittently. Eventually, the student spends long periods making music without receiving reinforcement from others because making music has itself become a reinforcer for doing that activity.

Some might explain the transition of our music student from an "externally reinforced person" to a "self-reinforced musician" as the development of intrinsic motivation, which seems to imply that something inside the person is responsible for maintaining the behavior. This view is incorrect from a behavioral standpoint. Applied behavior analysts describe intrinsic motivation as reinforcement that is received by manipulating the physical environment. Some individuals ride bicycles, go backpacking, read, write, or help others because manipulations of the environment provide reinforcement for engaging in those activities.

DEFINING BASIC INTERMITTENT SCHEDULES OF REINFORCEMENT

Ratio and Interval Schedules

Applied behavior analysts directly or indirectly embed ratio and interval intermittent schedules of reinforcement in most treatment programs, especially ratio schedules (Lattal & Neef, 1996). Ratio schedules require a number of responses before one response produces reinforcement. If the ratio requirement for a behavior is 10 correct responses, only the 10th correct response produces reinforcement. Interval schedules require an elapse of time before a response produces reinforcement. If the interval requirement is 5 minutes, reinforcement is contingent on the first correct response that occurs after 5 minutes has elapsed. Thereafter, reinforcement is contingent on the first correct response after 5 minutes has elapsed since the last reinforced response.

Ratio schedules require a number of responses to be emitted for reinforcement; an elapse of time does not change the number contingency. The participant's response rate determines the rate of reinforcement. The more quickly the person completes the ratio requirement, the sooner reinforcement will occur. Conversely, interval schedules require an elapse of time before a single response produces reinforcement. The total number of responses emitted on an interval schedule is irrelevant to when and how often the reinforcer will be delivered. Emitting a high rate of response during an interval schedule does not increase the rate of reinforcement. Reinforcement is contingent only on the occurrence of one response after the required time has elapsed. The availability of reinforcement is time-controlled with interval schedules, and rate of reinforcement is "self-controlled" with ratio schedules, meaning that the more quickly the individual completes the ratio requirement, the sooner reinforcement will occur.

Fixed and Variable Schedules

Applied behavior analysts can arrange ratio and interval schedules to deliver reinforcement as a fixed or a variable contingency. With a fixed schedule, the response ratio or the time requirement remains constant. With a variable schedule, the response ratio or the time requirement can change from one reinforced response to another. The combinations of ratio or interval and fixed or variable contingencies define the four basic schedules of intermittent reinforcement: fixed ratio, variable ratio, fixed interval, and variable interval.

SCHEDULE EFFECTS AND CONSISTENCY OF PERFORMANCE

The following sections define the four basic schedules of intermittent reinforcement, provide examples of each schedule, and present some well-established schedule effects derived from basic research.

Early experimental and applied behavior analysts (e.g., 1950–1970) reported similar consistencies of schedule effects among nonhuman animals and humans. Many behavior analysts today question the generality of the schedule effects presented in the next sections. This issue developed as an outcome of advancing research with human subjects, and the acceptance that humans—not nonhuman animals—have language. Verbal behavior allows humans to make and use rules that can influence their responses to schedules of reinforcement (i.e., rule-governed behavior). Jack Michael (1987) provides an example of a verbal rule controlling a schedule contingency.

> Humans differ behaviorally from nonhumans in one overwhelmingly significant way; humans have an extensive language repertoire which influences most aspects of their

> behavior. With the exception of very young children and the profoundly retarded the human subjects used in studies of the effects of complex contingencies have all had this language repertoire, and one would certainly expect it to play a role in determining the experimental results. One of my first attempts to study human behavior in a situation somewhat analogous to a nonhuman operant conditioning experiment involved a simple discrimination schedule. The human subject pressed a telegraph key and received—on a variable interval schedule—nickels as reinforcement. The VI schedule, however, was alternated with a period of extinction, and a light on the response panel was on when the VI schedule was in effect, and off when it was not. Only minimal instructions were given, and they did not mention the role of the light. My first subject pressed the telegraph key at a moderately high rate throughout the 30-minute session, and his rate during the light-off condition was no different than during the light-on condition. When he finished his session I asked him to tell me as much as he could about what he was doing. He stated that he received nickels for pressing the telegraph key—not for every press, but once in a while, and never when the light was off. I asked him why he continued to press when the light was off. He said he thought that maybe I was trying to see if he was a persistent type of person, one who would continue trying even though success seemed unlikely, and he wanted to show me that he was exactly that type of person.
>
> His explanation was quite reasonable, at least in its suggestion of the possible role of his social history and language repertoire, but I certainly did not interpret this event as a challenge to the generality of the effect of discrimination training on behavior. (pp. 38–39)

As you study the schedule effects sections that follow, recognize that participants' performance on schedules of reinforcement arranged in applied settings may differ considerably from that predicted by basic laboratory research as rule-governed behavior. Always remember, however, Michael's concluding comment on generality in the above quote. (See also "Perspectives on Using Schedules of Reinforcement in Applied Settings: Applied Research with Intermittent Schedules" in this chapter.)

Fixed Ratio Defined

A **fixed ratio (FR)** schedule of reinforcement requires the completion of a fixed number of responses for a reinforcer. For example, every fourth correct (or target) response produces reinforcement on an FR 4 schedule. An FR 15 schedule requires 15 responses to produce reinforcement. Skinner (1938) conceptualized each ratio requirement as a response unit. Accordingly, the response unit produces the reinforcer, not just the last response of the ratio.

Some business and industrial tasks are paid on an FR schedule (e.g., piecework). A worker might receive a pay credit after completing a specified number of tasks (e.g., assembling 15 pieces of equipment or picking a box of oranges). A student might receive either a happy face after learning 5 new sight words or a certain number of points after completing 10 math problems.

Jeffries, Crosland, and Miltenberger (2016) assessed the effectiveness of the tablet application *Look in My Eyes Steam Train* and differential reinforcement. During the tablet condition, a child looked into eyes on the iPad to see a number, and then picked a matching number on a grid. If the numbers matched, "a piece of coal appeared on the screen and a praise statement was provided by the [iPad] narrator." The interactive Steam Train game appeared on the tablet screen for 1 minute after the child received four pieces of coal. This example combined CRF and FR 4 schedules of reinforcement. Each correct number match produced a piece of coal and praise. Following four correct matches, the child received 1 minute of screen time to play the Steam Train game. (see also "Tandem Schedules" in this chapter under "Compound Schedules of Reinforcement").

Fixed Ratio Schedule Effects

Consistency of Performance

FR schedules produce a typical pattern of responding: (a) After the first response of the ratio requirement, the participant completes the required responses with little hesitation between responses; and (b) a **postreinforcement pause** follows reinforcement (i.e., the participant does not respond for a period following reinforcement). The *size of the ratio* and the *reinforcement magnitude* can influence the duration of the postreinforcement pause (Powell, 1969; Schlinger, Derenne, & Baron, 2008): Large ratio requirements may produce long pauses; small ratios produce short pauses. Lower reinforcement magnitudes—an inverse relation to the large ratio—may produce long pauses "indicating that magnitude and ratio size interact to determine the extent of pausing" (Schlinger et al., 2008, p. 43). Following the pause, the participant again completes the fixed ratio with little hesitation between responses. Analysts often identify this characteristic of FR responding as a *break-and-run* pattern.

Rate of Response

FR schedules often produce high rates of response. Quick responding on FR schedules maximizes the delivery of reinforcement because the higher the rate of response, the greater the rate of reinforcement. People work rapidly with a fixed ratio because they receive reinforcement upon completion of the ratio requirements.

The size of the ratio can influence the rate of response on FR schedules. To a degree, the larger the ratio requirement, the higher the rate of response. A teacher could reinforce every third correct answer to arithmetic facts. With this ratio requirement, the student might complete 12 problems within the specified time, producing reinforcement four times. The student might complete more problems in less time if the teacher arranged reinforcement contingent on 12 correct answers rather than 3. The higher ratio is likely to produce a higher rate of response. The rate of response decreases, however, if the ratio requirements are too large. The maximum ratio is determined in part by the participant's past FR history of reinforcement, motivating operations, the quality and magnitude of the reinforcer, and the

procedures that change the ratio requirements. For example, if ratio requirements are raised gradually over an extended period, extremely high ratio requirements can be reached.

Figure 13.1 depicts and summarizes the schedule effects typically produced by FR schedules of reinforcement.

Variable Ratio Defined

A **variable ratio (VR)** schedule of reinforcement requires the completion of a variable number of responses to produce a reinforcer. A number representing the average (e.g., mean) number of responses required for reinforcement identifies the VR schedule. For example, with a VR 10 schedule every 10th correct response on the average produces reinforcement. Reinforcement can come after 1 response, 14 responses, 5 responses, 19 responses, or *n* responses, but the average number of responses required for reinforcement is 10 (e.g., $1 + 14 + 5 + 19 + 11 = 50$; $50/5 = 10$).

The operation of a slot machine, the one-armed bandit, provides a good example of a VR schedule. These machines are programmed to pay off only a certain proportion of the times they are played. A player cannot predict when the next operation of the machine will pay off. The player might win 2 or 3 times in succession and then not win again for 20 or more plays.

De Luca and Holborn (1992) examined the effects of a VR schedule on three obese and three nonobese children's rate of pedaling an exercise bicycle. The children could use the exercise bicycle Monday to Friday during each week of the analysis, but received no encouragement to do so. The participants received the instruction to "exercise as long as you like" to initiate the baseline condition. De Luca and Holborn introduced the VR schedule of reinforcement after establishing a stable baseline rate of pedaling. They calculated the baseline mean number of pedal revolutions per minute and programmed the first VR contingency at approximately 15% faster pedaling than the baseline mean. The children received points on the VR schedule to exchange for backup reinforcers. De Luca and Holborn increased the VR schedule in two additional increments by approximately 15% per increment. All participants had systematic increases in their rate of pedaling with each VR value, meaning that the larger the variable ratio, the higher the rate of response. De Luca and Holborn reported that the VR schedule produced higher rates of response than did the FR schedule in their previous study (De Luca & Holborn, 1990). Figure 13.2 presents the participants' performances under baseline and VR (i.e., VR ranges of 70 to 85, 90 to 115, 100 to 130) conditions.

Student behaviors usually produce reinforcement following the completion of variable ratios. Normally a student cannot predict when the teacher will call on him to give an answer and thereby receive reinforcement. Good grades, awards, and promotions all may come after an unpredictable number of responses. And in checking seatwork, the teacher might reinforce a student's work after the completion of 10 tasks, another student's work after 3 tasks, and so on.

Variable Ratio Schedule Effects

Consistency of Performance

VR schedules produce consistent, steady rates of response. They typically do not produce a postreinforcement pause, as do FR schedules. Perhaps the absence of pauses in responding is due to the absence of information about when the next response will produce reinforcement. Responding remains steady because the next response may produce reinforcement.

Rate of Response

Like the FR schedule, the VR schedule tends to produce a quick rate of response. Also similar to the FR schedule, the size of the ratio influences the rate of response. To a degree, the larger the ratio requirement, the higher the rate of response. Again like FR schedules, when variable ratio requirements are thinned gradually over an extended period, participants will respond to extremely high ratio requirements. Figure 13.3 summarizes the schedule effects typically produced by VR schedules of reinforcement.

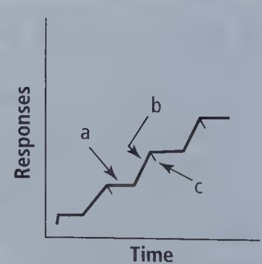
Variable Ratio Schedules in Applied Settings

Basic researchers use computers to select and program VR schedules of reinforcement. In applied settings, VR schedules are seldom implemented with a planned and systematic approach. In other words, the reinforcer is delivered by chance, hit or miss

Definition: Reinforcement delivered contingent on emission of a specified number of responses.

Schedule Effects:
After reinforcement a postreinforcement pause occurs. After the pause the ratio requirement is completed with a high rate of response and very little hesitation between responses. The size of the ratio influences both the pause and the rate.

Stylized Graphic Curve of Cumulative Responses:



a = postreinforcement pause
b = high rate of response "run"
c = reinforcer delivered upon emission of n^{th} response

Figure 13.1 Summary of FR schedule effects during ongoing reinforcement.

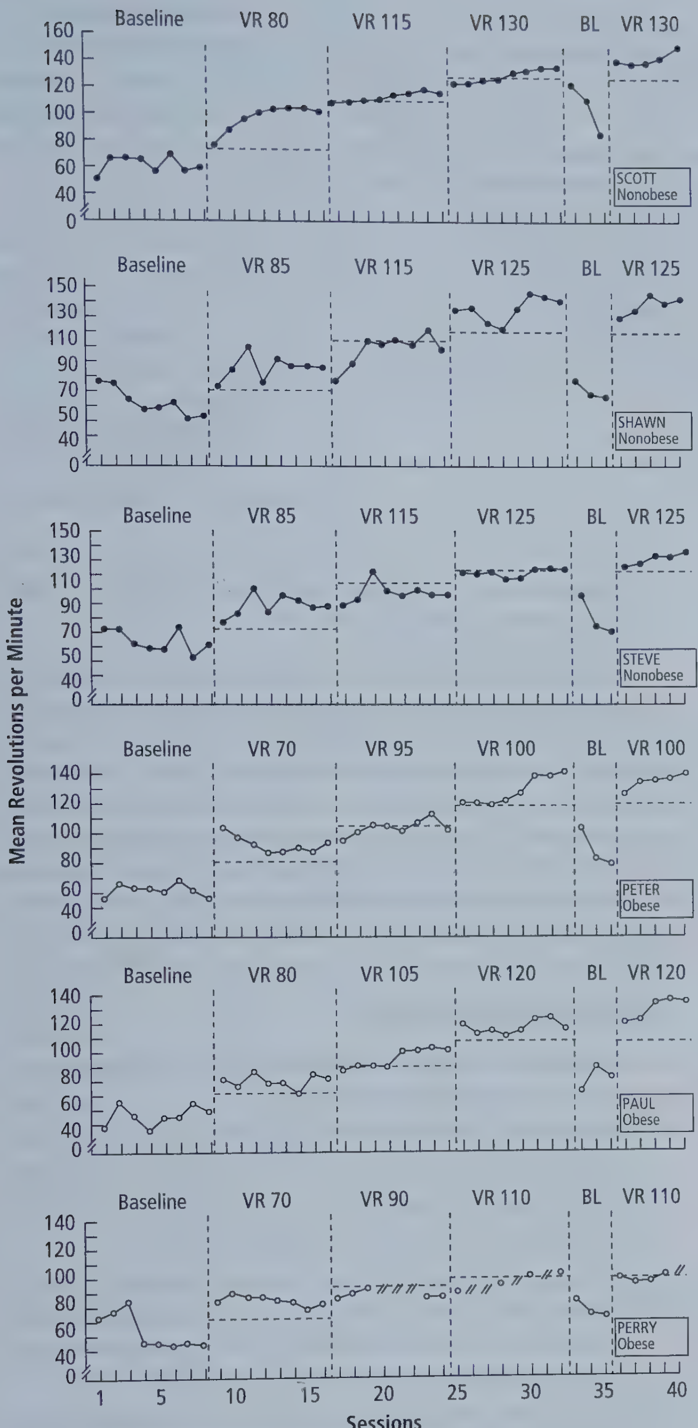


Figure 13.2 Mean revolutions per minute during baseline, VR 1 (VR range, 70 to 85), VR 2 (VR range, 90 to 115), and VR 3 (VR range, 100 to 130), return to baseline, and return to VR 3 phases for obese and nonobese subjects.

"Effects of a Variable-Ratio Reinforcement Schedule with Changing Criteria on Exercise in Obese and Nonobese Boys", R. V. De Luca and S. W. Holborn, 1992, Reproduced with permission of John Wiley and Sons Inc.

in most interventions. This nonsystematic delivery of reinforcement is not an effective use of VR schedules. Teachers can select and preplan VR schedules that approximate the VR schedules used in basic research. For example, teachers can plan variable ratios by (a) selecting a maximum ratio for a given activity (e.g., 15 responses) and (b) using a table of random numbers to produce the specific variable ratios for the schedule of reinforcement. A table of random numbers might generate the following sequence of ratios: 8, 1, 1, 14, 3, 10, 14, 15, and 6, producing a VR 8 schedule of reinforcement (on the average each eighth response produces the reinforcer) with the ratios ranging from 1 to 15 responses.

Teachers can apply the following VR procedures as individual or group contingencies of reinforcement for academic or social behavior:

Tic-Tac-Toe VR Procedure

1. The teacher establishes a maximum number for the individual student or group. The larger the maximum number selected, the greater the odds against meeting the contingency. For example, 1 chance out of 100 has less chance of being selected than 1 chance out of 20.
2. The teacher gives the individual or group a tic-tac-toe grid.

Definition: Reinforcer is delivered after the emission of a variable number of responses.

Schedule Effects: Ratio requirements are completed with a very high rate of response and little hesitation between responses. Postreinforcement pauses are not a characteristic of the VR schedule. Rate of response is influenced by the size of the ratio requirements.

Stylized Graphic Curve of Cumulative Responses:

a = high, steady rate of responding
b = reinforcement delivered after a varying number of required responses are emitted

Figure 13.3 Summary of VR schedule effects during ongoing reinforcement.

3. Students fill in each square of the grid with a number no greater than the maximum number. For example, if the maximum number is 30, the score sheet might look like this:

1	20	13
3	5	30
7	11	6

4. The teacher fills a box or some other type of container with numbered slips of paper (with numbers no higher than the maximum number). Each number should be included several times—for example, five 1s, five 2s, five 3s.
5. Contingent on the occurrence of the target behavior, students withdraw one slip of paper from the box. If the number on the paper corresponds to a number on the tic-tac-toe sheet, the students mark out that number on the grid.
6. The reinforcer is delivered when students have marked out three numbers in a row—horizontally, vertically, or diagonally.

For example, a student might withdraw one slip of paper for each homework assignment completed. Selecting an activity from the class job board (e.g., teacher's helper, collecting milk money, running the projector) could serve as the consequence for marking out three numbers in a row.

Classroom Lottery VR Procedure

1. Students write their names on index cards after successfully completing assigned tasks.
2. Students put signature cards into a box located on the teacher's desk.
3. After an established interval of time (e.g., 1 week), the teacher draws a signature card from the box and declares that student the winner. The lottery can have first, second, and third place, or any number of winners. The more cards students earn, the greater is the chance that one of their cards will be picked.

Teachers have used classroom lotteries with a variety of student accomplishments, such as reading non-assigned books. For example, for each book read, students write their names and the title of the book they have read on a card. Every 2 weeks the teacher picks one card from the box and gives the winning student a new book. To make the book an especially desirable consequence, the teacher lets students earn the privilege of returning the book to the school, inscribed with the student's name, class, and date (e.g., *Brian Lee, fifth grade, donated this book to the High Street Elementary School Library on May 8, 2019*).

Desk Calendar VR Procedure

1. Students receive desk calendars with loose-leaf date pages secured to the calendar base.
2. The teacher removes the loose-leaf date pages from the calendar base.
3. The teacher establishes a maximum ratio for the students.
4. The teacher numbers index cards consecutively from 1 to the maximum ratio. Multiple cards are included for each number (e.g., five 1s, five 2s). If a large average ratio is desired, the teacher includes more large numbers; for small average ratios, the teacher uses smaller numbers.
5. The teacher uses a paper punch to punch holes in the index cards for attaching the cards to the calendar base.
6. The teacher or student shuffles the index cards to quasi-randomize the order and attaches the index cards to a calendar base face down.
7. Students produce their own VR schedules by turning over one index card at a time. After meeting that ratio requirement, students flip the second card to produce the next ratio, and so on.

Students can use the desk calendar base to program VR schedules for most curriculum areas (e.g., arithmetic facts). For example, after receiving an arithmetic worksheet, the student flips the first card. It has a 5 written on it. After completing

five problems, she holds up her hand to signal her teacher that she has completed the ratio requirement. The teacher checks the student's answers, provides feedback, and presents the consequence for correct problems. The student flips the second card; the ratio requirement is 1. After completing that single problem, she receives another consequence and flips the third card. This time the ratio is 14. The cycle continues until all of the cards in the stack are used. New cards can then be added or old cards reshuffled to create a new sequence of numbers. The average of the numbers does not change in the reshuffling. (See Kauffman, Cullinan, Scranton, & Wallace, 1972, for a description of using a deck of playing cards, mounted on a refillable calendar base, to create FR and VR schedules of reinforcement that function in most environments.)

Fixed Interval Schedules

A **fixed interval (FI)** schedule of reinforcement provides reinforcement for the first response following a fixed duration of time. With an FI 3-min schedule, the first response following the elapse of 3 minutes produces the reinforcer. A common procedural misunderstanding with the FI schedule is to assume that the elapse of time alone is sufficient for the delivery of a reinforcer, assuming that the reinforcer is delivered at the end of each fixed interval of time. However, more time than the fixed interval can elapse between reinforced responses. The reinforcer is available after the fixed time interval has elapsed, and it remains available until the first response. When the first response occurs sometime after the elapse of a fixed interval, that response is immediately reinforced, and the timing of another fixed interval is usually started with the delivery of the reinforcer. This FI cycle is repeated until the end of the session.

Examples of FI schedules in everyday life are difficult to find. However, some situations do approximate and in reality function as FI schedules. For example, mail is often delivered close to a fixed time each day. An individual can make many trips to the mailbox to look for mail, but only the first trip to the mailbox following the mail delivery will produce reinforcement. Many textbook examples of FI schedules, such as the mail example, do not meet the definition of an FI schedule; but the examples do appear similar to an FI schedule. For example, receiving a paycheck as wages for work by the hour, day, week, or month is contingent on the first response on payday that produces the paycheck. Of course, earning the paycheck requires many responses during the interval that eventually lead to receiving the paycheck. In a true FI schedule, responses during the interval do not influence reinforcement.

FI schedules are relatively easy to use in applied settings. A teacher could make reinforcement available on an FI 2-min schedule for correct answers on an arithmetic worksheet. The teacher or student could use an electronic timer with a countdown function to signal the elapse of the 2-minute interval. The student's first correct answer following the interval produces reinforcement, and then the teacher resets the timer for another 2-minute interval. Similarly, the teacher could use a timing app or MotivAider® (www.habitchange.com) that vibrates to signal the elapse of an interval.

Fixed Interval Schedule Effects

Consistency of Performance

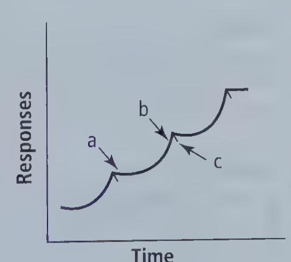
FI schedules typically produce a postreinforcement pause in responding during the early part of the interval. An initially slow but accelerating rate of response is evident toward the end of the interval, usually reaching a maximum rate just before delivery of the reinforcer. This gradually accelerating rate of response toward the end of the interval is called an *FI scallop* because of the rounded curves that are shown on a cumulative graph (see Figure 13.4).

FI postreinforcement pause and scallop effects can be seen in many everyday situations. When college students are assigned a term paper, they typically do not rush to the library and start to work on the paper immediately. More often they wait a few days or weeks before starting to work. However, as the due date approaches, their work on the assignment increases in

Definition: The first response after a designated and constant amount of time produces the reinforcer.

Schedule Effects: FI schedules generate slow to moderate rates of responding with a pause in responding following reinforcement. Responding accelerates toward end of the interval.

Stylized Graphic Curve of Cumulative Responses:



a = postreinforcement pause
b = increase in response rates as interval progresses and reinforcer becomes available
c = reinforcer delivered contingent on first correct response after interval

Figure 13.4 Summary of FI schedule effects during ongoing reinforcement.

an accelerating fashion, and many are typing the final draft just before class. Cramming for a midterm or final examination is another example of the FI scallop effect.

These examples with the reinforcement pause and scallop effects appear to be produced by FI schedules of reinforcement. They are not, however, because like the paycheck example, college students must complete many responses during the interval to produce the term paper or a good grade on the examinations, and the term paper and examinations have deadlines. With FI schedules, responses during the interval are irrelevant, and FI schedules have no deadlines for the response.

Why does an FI schedule produce a characteristic pause and scallop effect? After adjustment to an FI schedule, participants learn (a) to discriminate the elapse of time and (b) that responses emitted right after a reinforced response are never reinforced. Therefore, extinction during the early part of the interval might account for the postreinforcement pause. The effects of FI and FR schedules of reinforcement are similar in that both schedules produce postreinforcement pauses. However, it is important to recognize the different characteristics of behavior that emerge under each schedule. Responses under an FR schedule are emitted at a consistent rate until completing the ratio requirement, whereas responses under an FI schedule begin at a slow rate and accelerate toward the end of each interval.

Rate of Responding

Overall, FI schedules tend to produce a slow to moderate rate of response. The duration of the time interval influences the postreinforcement pause and the rate of response; to a degree, the larger the fixed interval requirement, the longer the postreinforcement pause and the lower the overall rate of response.

Variable Interval Schedules

A **variable interval (VI)** schedule of reinforcement provides reinforcement for the first correct response following the elapse of variable durations of time. The distinguishing feature of VI schedules is that "the intervals between reinforcement vary in a random or nearly random order" (Ferster & Skinner, 1957, p. 326). Behavior analysts use the average (i.e., mean) interval of time before the opportunity for reinforcement to describe VI schedules. For example, in a VI 5-min schedule the average duration of the time intervals between reinforcement and the opportunity for subsequent reinforcement is 5 minutes. The actual time intervals in a VI 5-min schedule might be 2 minutes, 5 minutes, 3 minutes, 10 minutes, or *n* minutes (or seconds).

Variable Interval Schedule Effects

Consistency of Performance

A VI schedule of reinforcement tends to produce a constant, stable rate of response. The slope of the VI schedule on a cumulative graph appears uniform with few pauses in responding (see Figure 13.5). A VI schedule typically produces few hesitations between responses. For example, pop quizzes at unpredictable times tend to occasion more consistent study behavior from students than do quizzes scheduled at fixed intervals of time. Furthermore, students are less apt to engage in competing off-task behaviors during instructional and study periods when a pop quiz is likely. The pop quiz is used often as an example of a VI schedule because the performance effect is similar to a VI performance. The pop quiz does not represent a true VI schedule, however, because of the required responses during the interval, and the deadline for receiving reinforcement.

Rate of Responding

VI schedules of reinforcement tend to produce low to moderate rates of response. Like the FI schedule, the average duration of the time intervals on VI schedules influences the rate of response; to a degree, the larger the average interval, the lower the overall rate of response. Figure 13.5 summarizes the schedule effects typically produced by VI schedules during ongoing reinforcement.

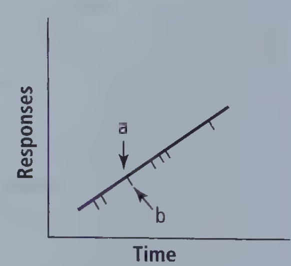
Variable Interval Schedules in Applied Settings

Basic researchers use computers to select and program VI schedules of reinforcement, as they do with VR schedules. Teachers seldom apply VI schedules in a planned and systematic way. For example, a teacher might set an electronic countdown timer with varied intervals of time ranging from 1 minute to 10 minutes

Definition: The first response following varying intervals of time produces the reinforcer.

Schedule Effects: VI schedules generate slow to moderate response rates that are constant and stable. There are few, if any, post-reinforcement pauses with VI schedules.

Stylized Graphic Curve of Cumulative Responses:



a = steady response rate; few, if any, postreinforcement pauses
b = reinforcer delivered

Figure 13.5 Summary of VI schedule effects during ongoing reinforcement.

without any prior plan as to which intervals or which order will be used. This set-them-as-you-go selection of intervals approximates the basic requirements for a VI schedule; however, it is not the most effective way of delivering reinforcement on a VI schedule. A planned, systematic application of varied intervals of time should increase the effectiveness of a VI schedule.

For example, applied behavior analysts can select the maximum time interval, whether in seconds or minutes, that will maintain performance and still be appropriate for the situation. Preferably, applied behavior analysts will use data from a direct assessment to guide the selection of the maximum VI interval, or at the least clinical judgment based on direct observation. Analysts can use a table of random numbers to select the varied intervals between 1 and the maximum interval, and then identify the VI schedule by calculating an average value for the VI schedule. The VI schedule may need adjustments following the selection of time intervals. For example, if a larger average interval of time appears reasonable, the teacher can replace some of the smaller intervals with larger ones. Conversely, if the average appears too large, the teachers can replace some of the higher intervals with smaller ones.

Interval Schedules with a Limited Hold

When a **limited hold** is added to an interval schedule, reinforcement remains available for a finite time following the elapse of the FI or VI interval. The participant will miss the opportunity to receive reinforcement if a targeted response does not occur within the time limit. For example, on an FI 5-min schedule with a limited hold of 30 seconds, the first correct response following the elapse of 5 minutes is reinforced, but only if the response occurs within 30 seconds after the end of the 5-minute interval. If no response occurs within 30 seconds, the opportunity for reinforcement has been lost and a new interval begins. The abbreviation LH identifies interval schedules using a limited hold (e.g., FI 5-min LH 30-sec, VI 3-min LH 1-min). Limited holds with interval schedules typically do not change the overall response characteristics of FI and VI schedules beyond a possible increase in rate of response.

Martens, Lochner, and Kelly (1992) used a VI schedule of social reinforcement to increase academic engagement of two 8-year-old boys in a third-grade classroom. The classroom teacher reported that the boys had serious off-task behaviors. The experimenter wore an earphone connected to a microcassette recorder that played a fixed-time cueing tape. During baseline, the cueing tape was programmed for a VI 20-sec schedule of reinforcement with a 0-sec limited hold for the availability of reinforcement. If the boys were not academically engaged when the 20-second interval timed out, they lost that opportunity for reinforcement until the next cue. Following baseline, the experimenter delivered contingent praise on VI 2-min or VI 5-min schedules of reinforcement that alternated daily on a quasi-random basis. Both boys' academic engagement on the VI 5-min schedule resembled their baseline engagement (see Figure 13.6). When the VI 2-min schedule was in effect, both students were academically engaged at consistently higher levels than they were during baseline or the VI 5-min condition.

Figure 13.7 summarizes and compares the four basic intermittent schedules of reinforcement.

THINNING INTERMITTENT REINFORCEMENT

Applied behavior analysts often use one of two procedures for **schedule thinning**. First, they thin an existing schedule by gradually increasing the response ratio or the duration of the time

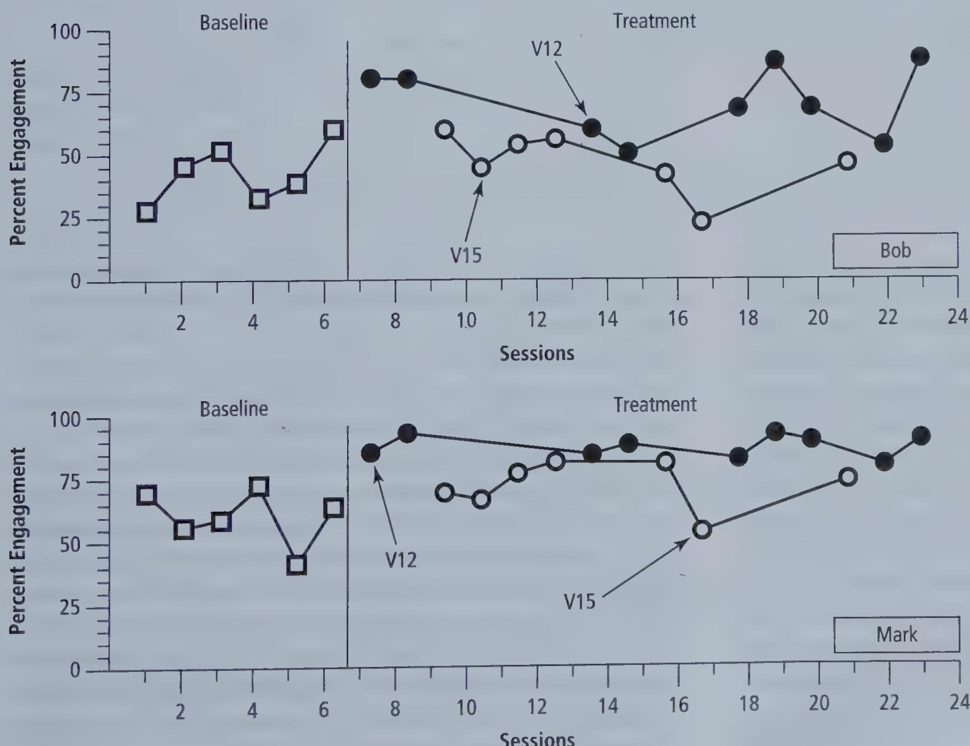
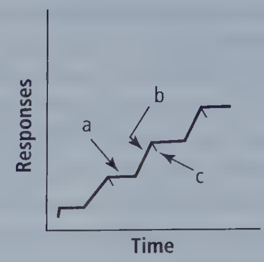
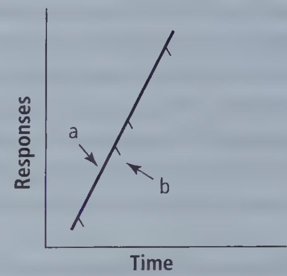
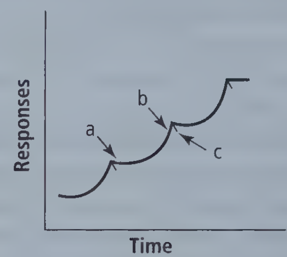
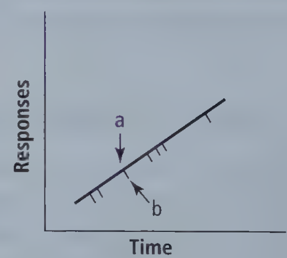


Figure 13.6 Academic engagement for two third-grade boys during baseline (VI 20-sec LH 0-sec) and treatment conditions with VI 2-min and VI 5-min schedules.

"The Effects of Variable-Interval Reinforcement on Academic Engagement: A Demonstration of Matching Theory", B. K. Martens, D. G. Lochner, and S. Q. Kelly, 1992, Reproduced with permission of John Wiley and Sons Inc.

Figure 13.7 Four basic intermittent schedules of reinforcement compared.

Schedule	Schedule Effects	Graphic Depiction
Fixed ratio (FR)—reinforcement follows a fixed number of responses (c)	• High overall rate of responding • Postreinforcement pause (a) • Then a *break-and-run* pattern of responding with little hesitation between responses (b) • Maximum ratio affected by past FR history of reinforcement, motivating operations, reinforcer quality, and changes in ratio requirements • Abrupt, large increases in FR size may generate ratio strain (avoidance, aggression, unpredictable pauses in responding)	
Variable ratio (VR)—reinforcement follows a variable number of responses (b)	• High, steady rate of responding (a) • No postreinforcement pause • High VR response requirements can be attained by gradually thinning the schedule	
Fixed interval (FI)—reinforcement follows the first response after a fixed amount of time has elapsed since the previous reinforced response (c)	• Low to moderate overall rate of responding • Postreinforcement pause (a) • Response rate accelerates (b) as end of the interval approaches, producing a scallop effect • Longer FI intervals produce the longer postreinforcement pauses and lower overall response rates	
Variable interval (VI)—reinforcement follows the first response after a variable amount of time has elapsed since the previous reinforced response (b)	• Consistent, stable rate of responding (a) • No postreinforcement pause • Low to moderate rates of responding • Longer mean intervals produce lower overall response rates	

interval. If a student has answered addition facts effectively and responded well to a CRF schedule for two or three sessions, the teacher might thin the reinforcement contingency slowly from one correct addition fact (CRF) to a VR 2 or VR 3 schedule. The student's performance should guide the progression from a dense schedule (i.e., responses produce frequent reinforcement) to a thin schedule (i.e., responses produce less frequent reinforcement). Applied behavior analysts should use small increments of schedule changes during thinning and ongoing evaluation of the learner's performance to adjust the thinning process and avoid the loss of previous improvements.

Second, teachers often use instructions to clearly communicate the schedule of reinforcement, facilitating a smooth transition during the thinning process. Instructions include rules, directions, and signs. Participants do not require an awareness of environmental contingencies for effective intermittent reinforcement, but instructions may enhance the effectiveness of interventions when participants are told what performances produce reinforcement.

Ratio strain can result from abrupt increases in ratio requirements when moving from denser to thinner reinforcement schedules. Also, ratio strain will occur when the ratio becomes so large that the reinforcement cannot maintain the response level, or the response requirement exceeds the participant's physiological capabilities. Common behavioral characteristics associated with ratio strain include avoidance, aggression, and unpredictable pauses in responding. Applied behavior analysts should reduce the ratio requirement when ratio strain is evident. The analyst can again gradually thin ratio requirements after recovering the behavior. Preemptively, small and gradual increases in ratio requirements help to avoid the development of ratio strain.

VARIATIONS ON BASIC INTERMITTENT SCHEDULES OF REINFORCEMENT

Schedules of Differential Reinforcement of Rates of Responding

Applied behavior analysts frequently encounter behavior problems that result from the rate that people perform certain behaviors. Responding too infrequently, or too often, may be detrimental to social interactions or academic learning. Differential reinforcement provides an intervention for behavior problems associated with rate of response. Differential reinforcement of particular rates of behavior is a variation of ratio schedules. Delivery of the reinforcer is contingent on responses occurring at a rate either higher than or lower than some predetermined criterion. The reinforcement of responses higher than a predetermined criterion is called **differential reinforcement of high rates (DRH)**. When responses are reinforced only when they are lower than the criterion, the schedule provides **differential reinforcement of low rates (DRL)**. DRH schedules produce a higher rate of responding. DRL schedules produce a lower rate of responding.

Applied behavior analysts use three definitions of DRH and DRL schedules. The first definition states that reinforcement is available only for responses that are separated by a given duration of time. This first definition is sometimes called spaced-responding DRH or spaced-responding DRL. An interresponse time (IRT) identifies the duration of time that occurs between two responses. IRT and rate of response are functionally related. Long IRTs produce low rates of responding; short IRTs produce high rates of responding. Responding on a DRH schedule produces reinforcement whenever a response occurs before a time criterion has elapsed. If the time criterion is 30 seconds, the participant's response produces reinforcement only when the IRT is 30 seconds or less.

Under the DRL schedule, a response produces reinforcement when it occurs after a time criterion has elapsed. If the stated DRL time criterion is again 30 seconds, a response produces reinforcement only when the IRT is 30 seconds or greater.

This first definition of DRH and DRL as IRT schedules of reinforcement has been used almost exclusively in laboratory settings. There are two apparent reasons for its lack of application in applied settings: (a) Most applied settings do not have sufficient automated equipment to measure IRT and to deliver reinforcement using an IRT criterion; and (b) reinforcement is delivered usually, but not necessarily, following each response that meets the IRT criterion. Such frequent reinforcement would disrupt student activity in most instructional settings. However, with greater use of computers for tutorial and academic response practice, opportunities should increase for using IRT-based schedules of reinforcement to accelerate or decelerate academic responding. Computers can monitor the pauses between academic responses and provide consequences for each response meeting the IRT criterion, with little disruption in instructional activity.

Based on the laboratory procedures for programming DRL schedules presented previously, Deitz (1977) labeled and described two additional procedures for using differential reinforcement of rates of responding in applied settings: full-session DRH or DRL and interval DRH or DRL. Deitz initially used the full-session and interval procedures as a DRL intervention for problem behaviors. The full-session and interval procedures, however, apply also for DRH.

A DRH full-session schedule provides reinforcement if the total number of responses during the session meets or exceeds a number criterion. If the participant emits fewer than the specified number of responses during the session, the behavior is not reinforced. The DRL full-session schedule is procedurally the same as the DRH schedule, except reinforcement is provided for responding at or below the criterion limit. If the participant emits more than the specified number of responses during the session, reinforcement is not delivered.

The interval definition for DRH and DRL schedules states that reinforcement is available only for responses that occur at a minimum or better rate of response over short durations of time during the session. To apply an interval DRH schedule, the analyst organizes the instructional session into equal intervals of time and dispenses a reinforcer at the end of each interval when the student emits a number of responses equal to, or greater than, a number criterion. The interval DRL schedule is procedurally like the DRH interval schedule, except that reinforcement is provided for responding at or below the criterion limit.

The **differential reinforcement of diminishing rates (DRD)** schedule provides reinforcement at the end of a predetermined time interval when the number of responses is less than a criterion that is gradually decreased across time intervals based on the individual's performance (e.g., fewer than five responses per 5 minutes, fewer than four responses per 5 minutes, fewer than three responses per 5 minutes). Deitz and Repp (1973) used a group DRD contingency to reduce off-task talking of 15 high school senior girls. They set the first DRD criterion limit at five or fewer occurrences of off-task talking during each 50-minute class session. The DRL criterion limits were then gradually reduced to three or fewer, one or fewer, and finally no responses. The students earned a free Friday class when they kept off-task talking at or below the DRD limit Monday through Thursday.

The previous example of a DRD schedule used a procedure identical to that described for the full-session DRL. DRD is also a procedural variation on interval DRL schedules described by Deitz (1977) and Deitz and Repp (1983). The typical procedure for using an interval DRL as an intervention for problem behavior provided reinforcement contingent on emitting one or no responses per brief interval. After the problem behavior stabilizes at the initial criterion, the analyst maintains the maximum criterion of one or no responses per interval, but increases the duration of the session intervals to further diminish the behavior. Increasing the duration of session intervals continues gradually until the problem behavior achieves a terminal low rate of responding.

Deitz and Repp (1983) programmed the interval DRL with a criterion greater than one response per interval, then gradually diminished the maximum number of responses per interval while the duration of the interval remained constant (e.g., fewer than five responses per 5 minutes, fewer than four responses per 5 minutes, fewer than three responses per 5 minutes, etc.). The DRD schedule and the interval DRL schedule that use a

maximum number criterion greater than one per interval are different terms for the same procedure. Full-session and interval DRL have a long history of application in applied behavior analysis. DRD offers applied behavior analysts a new, and perhaps improved, label for the interval DRL procedure. (For a detailed description of differential reinforcement of rates of response, see Chapter 25.)

Increasing Response Variability with Lag Schedules of Reinforcement

Response variability often sets the occasion for expanding important behavior repertoires such as developing problem-solving skills, improving language acquisition and communication behaviors, and facilitating enhanced academic and social behaviors. Behavior analysts consider response variability an operant and therefore sensitive to reinforcement contingencies (Neuringer, 2004; Page & Neuringer, 1985). A **lag schedule** of reinforcement is one method to increase response variability. Reinforcement on a lag schedule is contingent on a response differing in some predetermined way (e.g., different topography, different sequence) from one or more responses that preceded it. With a Lag 1 schedule, each response that differs from the prior response produces reinforcement. Reinforcement on a Lag 2 schedule is contingent on a response being different from the previous 2 responses, a Lag 3 requires a response to differ from the previous 3 responses, and so on. To produce reinforcement on a lag infinity schedule, a response must differ from all prior responses.

Wiskow and Donaldson (2016) used a lag schedule to promote varied naming of items in the categories of vegetable and animal. Three children demonstrating low levels of variability during initial testing served as participants. The children included two typically developing girls 3 and 4 years old, and a 6-year-old boy diagnosed with high-functioning autism.

> Lag 0. The therapist gave the initial instruction, "We are going to play the category game and see how many things you can think of. Wait quietly until it is your turn. When you tell me a correct answer, you will earn a token. Let's start with [category]." The therapist provided general feedback following an incorrect answer (e.g., "Nice try, but a banana is not a vegetable."). The therapist delivered specific praise (e.g., "Great! Lettuce is a vegetable!") and a token following correct responses. If the child did not respond, the therapist said nothing and moved on to the next child.
>
> Lag 1. We arranged a Lag 1 schedule only in the animal category. The therapist gave the following initial instruction in Lag 1 sessions: "We are going to play the category game and see how many things you can think of. Wait quietly until it is your turn. When you tell me an answer that is different than what someone else said, you will earn a token. Let's start with [category]." We did not include the first correct response in the lag schedule because this schedule required that a child's response be different than the immediately prior correct response. That is, in the Lag 1 condition, the first participant could say any correct response and receive praise and a token. Incorrect responses did not reset the schedule or factor into the schedule requirement and were not included when we calculated the percentage of variability (pp. 476–477).

Behavior analysts have increasingly addressed the importance of response variability in the development of verbal behavior (e.g., *intraverbals*, Contreras & Betz, 2016; *tacts*, Heldt & Schlinger, Jr., 2012; *phonemic variability*, Koehler-Platten, Grow, Schulze, & Bertone, 2013; *echoics*, Esch–J.W., Esch–B.E., & Love, 2009; [see also Chapter 18]). These researchers (and others) have demonstrated the effectiveness of Lag schedules on increasing variability with different behaviors, including verbal behavior and with other different categories of participants (e.g., age, gender, disability).

Heldt and Schlinger (2012) investigated the effects of a Lag 3 schedule on increasing variability and, importantly, on maintaining the increased variability. A 13-year-old boy with autism and fragile X syndrome and a 4-year-old boy with mild intellectual disability served as participants. During the Lag 3 schedule, the instructor began the 10 trials per session by (a) saying the participant's name, (b) presenting a picture with an array of images, and (c) saying "What do you see?" Reinforcement was provided when the fourth response differed from the previous three responses. "Therefore, for the first three responses, instructors utilized a reinforcement of different forms procedure, such as the one described by Goetz and Baer (1973). By the fourth response, the instructor implemented the Lag 3 schedule regardless of the prior four responses (i.e., whether they were all novel, all rote, all irrelevant responses, or some combination of those). On this schedule, the instructor ignored incorrect responses, no response within 5 s, or responses that did not differ from the previous three responses, and proceeded to the next trial (p. 133)." The follow-up occurred 3 weeks after termination of the Lag 3 sessions. The results showed that both participants increased variability in tacting during the Lag 3 sessions, and that variability in responding was maintained during a 3-week follow-up. (See also Lag schedule, Chapter 30.)

Progressive Schedules of Reinforcement

A **progressive schedule of reinforcement** systematically thins each successive reinforcement opportunity within the session independent of the participant's behavior. Progressive-ratio and delay to reinforcement (sometimes called *progressive interval, time-based progressive delay, delay fading*) schedules of reinforcement are used for reinforcer assessment and behavioral intervention.

Progressive-ratio Schedule of Reinforcement

In 1961, Hodos introduced a variation of the fixed ratio (FR) schedule of reinforcement that increased the ratio requirements in incremental steps within the session. For example, the participant begins the session on an FR 5 schedule of reinforcement, and receives reinforcement after emitting five responses. Following FR 5 responses and reinforcement, the schedule of reinforcement is now increased by some specified amount (e.g., from FR 5 to FR 10). Following the FR 10 responses

and reinforcement, the ratio would again increase (e.g., from FR 10 to FR 15). This pattern of increasing ratio requirements continues to the session end. Analysts change PR schedule requirements using (a) arithmetic progressions to add a constant number to each successive ratio or (b) geometric progressions to add successively a constant proportion of the preceding ratio (Lattal & Neef, 1996). These within-session changes define the **progressive-ratio (PR) schedule of reinforcement** (Poling, 2010; Roane, 2008; Schlinger et al., 2008).[1,2]

During the PR sessions, the ratios are typically increased to the "breaking point," when the participant stops responding. Comparing the breaking points and corresponding number of responses associated with each PR condition offers an assessment for the potency of a reinforcer. For example, Jerome and Sturmey (2008) had three adults with developmental disabilities complete verbal and pictorial assessments to identify staff preferences. Following the verbal and pictorial assessments, Jerome and Sturmey used a PR schedule to evaluate the reinforcer potency of interactions with preferred staff. The breaking points with the PR schedules were higher for all three adults under responses for positive social interactions with their preferred staff than with their nonpreferred staff.

> In addition to the use of break points as a measure of reinforcer efficacy, PR schedules are also characterized by the omission of a terminal schedule value. That is, PR schedule requirements typically increase throughout the course of a session until responding stops for a period of time (e.g., 5 min) or until a predetermined duration of the observation has been reached (a session cap). Thus, total response output (as opposed to response rate) and break-point values are the primary measures of interest when evaluating behavior with PR schedules. (Roane, 2008, p. 155)

In an article on the applied use of PR schedules, Roane (2008) recommended several areas for future research, including

- The initial ratio size
- The progression of the step-size ratios
- The influence of response topographies on PR schedule performance
- The likely interaction of increasing ratios and the values of reinforcer delivery
- Using PR schedules to evaluate the therapeutic effects of psychopharmacological treatments
- The use of translational research with PR schedules, linking applied and basic research

Words of Caution When Using PR Schedules

Poling (2010) identified the following concerns regarding PR schedules with protected populations and people with special needs.

> In applied research—at least, when *applied* is defined as proposed in Baer, Wolf, and Risley's (1968) seminal article—the potency of a scheduled reinforcer is important primarily with respect to whether or not that reinforcer can be arranged to improve a socially significant target behavior. How it affects behavior in other circumstances, even in the population of concern (e.g., children with autism), is of value only if (a) this information is easily obtained, so there is little cost for people with special needs; and (b) this information allows applied behavior analysts to predict the clinical utility of the scheduled reinforcer. PR schedules are not an especially quick way to scale reinforcer potency. In addition, exposure to long ratios (as when the breaking point is approached) is known to be aversive (e.g., Dardano, 1973). Therefore, members of protected populations should not be exposed to them unless there are clear offsetting benefits. To date, no such benefits have been demonstrated, and no compelling arguments for their existence have been provided. (p. 349).

Delay to Reinforcement Schedules

Applied behavior analysts have used delay to reinforcement schedules to develop self-control (e.g., Binder, Dixon, & Ghezzi, 2000; Dixon & Cummins, 2001). For example, Dixon and Holcomb (2000) used a delay schedule to develop cooperative work behaviors and self-control of six adults dually diagnosed with intellectual disabilities and psychiatric disorders. The adults participated in two groups composed of three men in Group 1 and three women in Group 2. During a natural baseline condition, the groups received instruction to exchange or share cards to complete a cooperative task of sorting playing cards into piles by categories (e.g., hearts with hearts). Dixon and Holcomb terminated a natural baseline session for the group when one of the adults stopped sorting cards.

The groups received points for working on the card-sorting task during the choice baseline condition and the self-control training condition. Groups exchanged their points earned for items such as soda pop or cassette players, ranging in values from 3 points to 100 points.

During the choice baseline conditions, the group's participants could choose an immediate 3 points before doing the card sorting or a delayed 6 points after sorting the cards. Both groups chose the immediate smaller number of points rather than the larger amount following a delay to reinforcement.

During self-control training, the participants were asked while working on a cooperative task, "Do you want 3 points now, or would you like 6 points after sorting the cards for Z minutes and seconds?" (pp. 612–613). The delay was initially 0 seconds for both groups. The progressive delay to reinforcement ranged from an increase of 60 seconds to 90 seconds following each session that the group performance met the exact criterion for number of seconds of task engagement. The terminal goals for the delay to reinforcement were 490 seconds for Group 1 and 772 seconds for Group 2. Both groups achieved these delay-to-reinforcement goals. Following the introduction of the progressive delay procedure, both groups improved their cooperative work engagement and the self-control necessary to select progressively larger delays to reinforcement that resulted in more points earned. Figure 13.8 shows the performance of both groups of adults during natural baselines, choice baselines, and self-control training conditions.

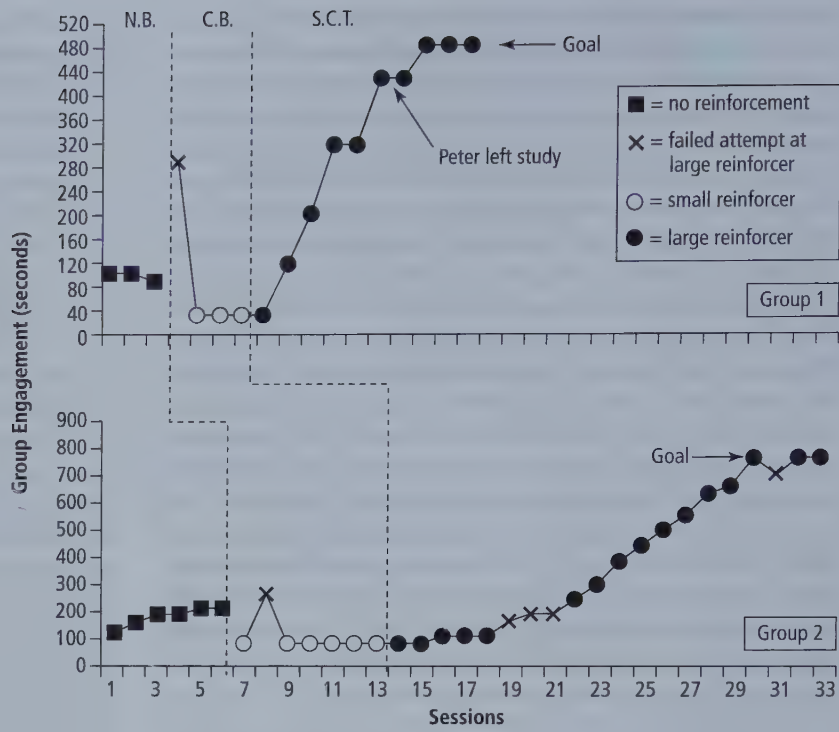


Figure 13.8 Number of seconds of engagement in the concurrent delay activity of cooperative card sorting during natural baseline (N.B.), choice baseline (C.B.), and self-control training (S.C.T.) for each group of participants. Filled circles represent performance at exactly the criterion level, and X data points represent the number of seconds of engagement below the criterion.

From "Teaching Self-Control to Small Groups of Dually Diagnosed Adults," by M. R. Dixon and S. Holcomb, 2000, *Journal of Applied Behavior Analysis, 33*, p. 613. Copyright 1992 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

Applied behavior analysts often use functional communication training (FCT)—the most prescribed function-based intervention to reduce problem behavior (e.g., aggression, stereotypy, self-injurious behavior [SIB], making threats, elopement) (Greer, Fisher, Saini, Owen, & Jones, 2016). Resurgence may occur when FCT is discontinued (see Chapter 24), or with a rapid thinning of the schedule of reinforcement. If resurgence of a previously reduced problem behavior occurs, applied behavior analysts might apply a delay to reinforcement schedule.

For example, Stevenson, Ghezzi, & Valenton (2016) used FCT (Carr & Durand, 1985) and a delay to reinforcement schedule (called delay fading) to treat the elopement behavior of a 9-year-old boy with autism. The FCT began with a functional behavior assessment (see Chapter 27)—(a) to identify reinforcers and (b) to identify processes for accessing reinforcement with appropriate behavior. During FCT, the boy learned to ask appropriately for the identified reinforcers and his eloping decreased. Stevenson et al. applied the delay fading condition following the boy's behavioral improvements during the FCT condition. The delay fading method used a conjunctive FR 1/VI schedule of reinforcement.

> Throughout these sessions, when Damon [the child] asked for an item or activity (i.e., the FR 1 schedule), the tutor said, "You have to wait for now, so let's keep walking." The VI schedule began timing at this moment, and the first appropriate request that Damon made after the interval elapsed produced the relevant item or activity. The value of the VI component progressed through various phases, when no elopement occurred for three consecutive sessions, from VI 15 s to VI 30 s, VI 45 s, VI 90 s, and finally, VI 300 s. (Stevenson et al., 2016, p. 171)

The results from the FCT and delay fading conditions indicated that the child continued with enduring appropriate communication and low elopements. He learned to wait for desired activity. (See Conjunctive Schedules later in this chapter).

COMPOUND SCHEDULES OF REINFORCEMENT

Applied behavior analysts combine two or more basic schedules of reinforcement—continuous reinforcement (CRF), the four intermittent schedules of reinforcement (FR, VR, FI, VI), differential reinforcement of various rates of responding (DRH, DRL), and extinction (EXT)—to form **compound schedules of reinforcement**. The basic component schedules:

- can occur simultaneously or successively;
- can occur with or without discriminative stimuli; and
- make reinforcement contingent on meeting the response requirements of each component schedule independently, or contingent on meeting the response requirements of the combination of the compound schedules (Ferster & Skinner, 1957).

Concurrent Schedules

A **concurrent schedule (conc)** of reinforcement occurs when (a) two or more contingencies of reinforcement (b) operate independently and simultaneously (c) for two or more behaviors. Each schedule is correlated with a discriminative stimulus. People in the natural environment have opportunities for making choices among concurrently available events. For example, Sharon receives a weekly allowance from her parents, contingent on doing daily homework and cello practice. After school she can choose when to do homework and when to practice the cello, and

she can distribute her responses between these two simultaneously available schedules of reinforcement. Applied behavior analysts use concurrent schedules for reinforcer assessment and for behavioral interventions.

Using Concurrent Schedules for Reinforcer Assessment

Applied behavior analysts have used concurrent schedules extensively to provide choices during the assessment of consequence preferences and the assessment of response quantities (e.g., rate, latency, magnitude) and reinforcer quantities (e.g., rate, duration, immediacy, amount). Responding to concurrent schedules provides a desirable assessment procedure because (a) the participant makes choices, (b) making choices during assessment approximates the natural environment, (c) the schedule is effective in producing hypotheses about potential reinforcers operating in the participant's environment, and (d) these assessments require the participant to choose between stimuli rather than indicating a preference for a given stimulus (Adelinis, Piazza, & Goh, 2001; Neef, Bicard, & Endo, 2001; Piazza et al., 1999).

Roane, Vollmer, Ringdahl, and Marcus (1998) presented 10 items to a participant, 2 items at a time. The participant had 5 seconds to select 1 item by using a reaching response to touch the selected item. As a consequence for the selection, the participant received the item for 20 seconds. The analyst verbally prompted a response if the participant did not respond within 5 seconds, waiting another 5 seconds for the occurrence of a prompted response. Items were eliminated from the assessment (a) if they were not chosen during the first five presentations or (b) if they were chosen two or fewer times during the first seven presentations. The participant made a total of 10 choices among the remaining items. The number of selections out of the 10 opportunities served as a preference index. (See also "Concurrent Schedule Reinforcer Assessment" in Chapter 11.)

Using Concurrent Schedules for Intervention

Applied behavior analysts have used concurrent schedules extensively for improving vocational, academic, and social skills in applied settings (e.g., Cuvo, Lerch, Leurquin, Gaffaney, & Poppen, 1998; Reid, Parsons, Green, & Browning, 2001; Romaniuk et al., 2002). For example, Hoch, McComas, Johnson, Faranda, and Guenther (2002) arranged two concurrent response alternatives for three boys with autism. The boys could play in one setting with a peer or sibling, or play alone in another area. Hoch and colleagues manipulated the duration of access to toys (i.e., reinforcer magnitude) and preference (i.e., reinforcer quality). In one condition, the magnitude and quality of the reinforcer were equal in both settings. In the other condition, the magnitude and quality of the reinforcer were greater for play in the setting with a peer or sibling than in the play-alone setting. With the introduction of the condition with greater magnitude and quality of the reinforcer, the boys allocated more play responses to the setting with the peer or sibling, rather than playing alone. The magnitude and quality of the reinforcer influenced choices made by the three boys. Figure 13.9 reports the percentage of responses allocated to the concurrent play areas.

Concurrent Schedule Performances: Formalizing the Matching Law

Cuvo et al. (1998) reported that concurrent schedules typically produce two response patterns. With concurrent interval schedules (conc VI/VI, conc FI/FI), participants "typically do not allocate all of their responses exclusively to the richer schedule [i.e., the schedule producing the higher rate of reinforcement]; rather, they distribute their responding between the two schedules to match or approximate the proportion of reinforcement that is actually obtained on each independent schedule" (p. 43). Conversely, with concurrent ratio schedules (conc VR/VR, conc FR/FR), participants are sensitive to the ratio schedules and tend to maximize reinforcement by responding primarily to the ratio that produces the higher rate of reinforcement.

Williams (1973) identified three types of interactions found with concurrent schedules. First, when similar reinforcement is scheduled for each of the concurrent responses, the response receiving the higher frequency of reinforcement will increase in rate, whereas a corresponding decrease will occur in the response rate of the other behavior. Second, when one response produces reinforcement and the other produces punishment, responses associated with punishment will decrease in occurrence. That decrease may produce a higher rate of response for the behavior producing reinforcement. Third, with a concurrent schedule programmed for one response to produce reinforcement and the other response to produce avoidance of an aversive stimulus, the rate of avoidance responding will accelerate with an increase in the intensity or the frequency of the aversive stimulus. As avoidance responding accelerates, typically responding on the reinforcement schedule will then decrease.

The **matching law** addresses response allocation to choices available with concurrent schedules of reinforcement. Typically, the rate of responding is proportional to the rate of reinforcement received from each choice alternative. The characteristics of performance on concurrent schedules are consistent with the relationships formalized by Herrnstein (1961, 1970) and as detailed by Cuvo et al. (1998) and Williams (1973).

Discriminated Compound Schedules of Reinforcement

A discriminative stimulus (S^D) signals the presence of each component schedule in a discriminated compound schedule of reinforcement.

Multiple Schedules

A **multiple schedule (mult)** presents two or more basic schedules of reinforcement in an alternating, usually random, sequence. The basic schedules within the multiple schedule occur successively and independently. A discriminative stimulus is correlated with each basic schedule, and that stimulus is present as long as the schedule is in effect. Applied behavior analysts have used multiple schedules of reinforcement for diverse treatment conditions—for example, thinning dense schedules of reinforcement (Greer, Fisher, Saini, Owen, & Jones, 2016), facilitating the verbal skill of children with autism (e.g., Sidener, Shabani, Carr, &

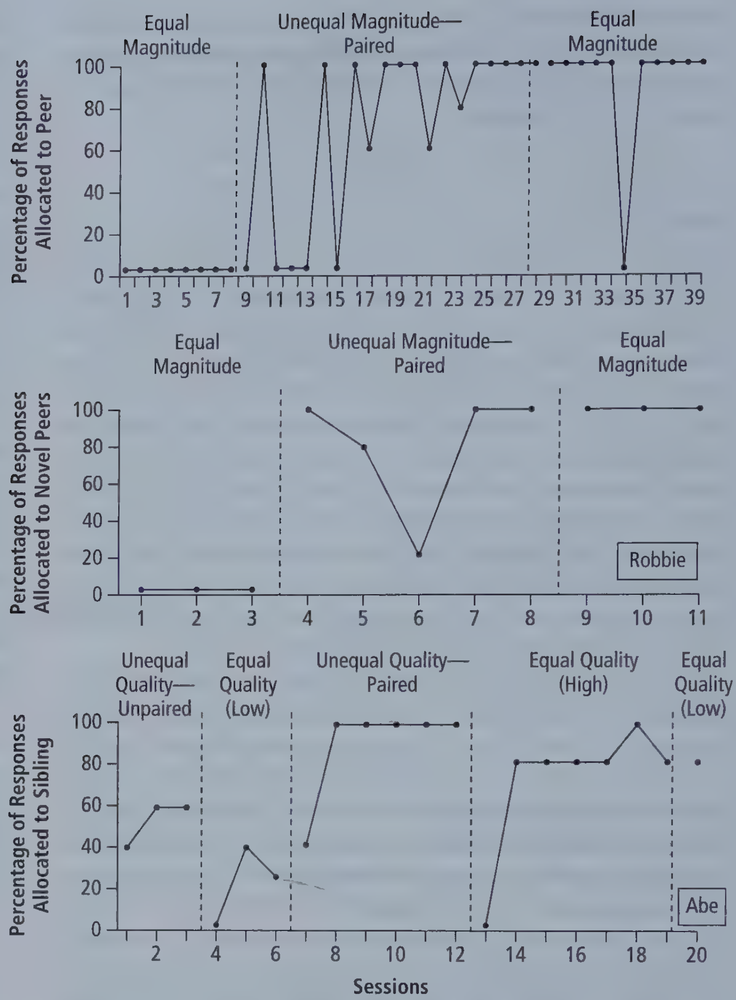


Figure 13.9 Percentage of responses allocated to the play area with the peer across experimental sessions (top panel) and in natural-setting probes with different peers in the classroom (middle panel) for analyzing the magnitude of reinforcement with Robbie, and the percentage of responses allocated to the play area with the sibling across experimental sessions for analyzing the quality of reinforcement with Abe (bottom panel).

"The Effects of Magnitude and Quality of Reinforcement on Choice Responding During Play Activities" by H. Hoch, J. J. McComas, L. Johnson, N. Faranda, and S. L. Guenther, 2002, Journal of Applied Behavior Analysis, Reproduced with permission of John Wiley and Sons Inc.

Roland, 2006), and reducing severe problem behavior (Neidert, Iwata, & Dozier, 2005). Academic behaviors also can become sensitive to the control of multiple schedules of reinforcement. For example, a student might respond to basic arithmetic facts with her teacher, and also with her tutor. With the teacher, the student responds to arithmetic facts during small-group instruction. The tutor then provides individual instruction and practice on the facts. This situation follows a multiple schedule because there is one class of behavior (i.e., math facts), a discriminative stimulus for each contingency in effect (i.e., teacher/tutor, small group/individual), and different conditions for reinforcement (i.e., reinforcement is less frequent in group instruction).

Tiger and Hanley (2005) used a three-component multiple schedule of reinforcement ($SR+/EXT_1/EXT_2$) to teach two preschool children how and when to recruit teacher attention at appropriate times. The children sat at individual tables with access to materials such as blocks and string beads. The analyst looked down except for giving contingent attention to one child's approach responses (SR+), while the other child's approach experienced extinction (EXT_1), and both children's approaching the analyst produced extinction (EXT_2). The analyst wore a red lei (discriminative stimulus) when attention was available for one child, a blue lei when attention was available to the other child, and a white lei when attention was not available to either child. "The purpose of the MULT condition was to determine if correlating stimuli with the availability (SR+) and unavailability (EXT_1 and EXT_2) of attention was sufficient to bring children's social approaches under stimulus control." (p. 502) The children demonstrated discriminated performances when—prior to the session—they received verbal reminders of the multiple schedule arrangements (cf. Tiger & Hanley, 2004; Tiger, Hanley, & Heal, 2006).

Torelli, Lloyd, Diekman, and Wehby (2016) extended Tiger and colleagues' (2004, 2005, 2006) line of research from preschool to elementary school classrooms. In the words of Torelli and colleagues, "The purpose of the current study was to evaluate the effects of a class-wide multiple schedule on differentiated rates of student recruitment of teacher attention in two public elementary classrooms. General education teachers implemented the multiple schedule intervention in the context of a common instructional routine (i.e., small group rotations among reading centers). Results indicated that the multiple schedule intervention was effective at decreasing disruptive bids for attention when teacher attention was not available" (p. 1).

To learn more about the use of multiple schedules in clinical assessment and treatment, see Saini, Miller, and Fisher (2016) and "Multiple Schedule Reinforcer Assessment" in Chapter 11.

Reynolds (1961) introduced the term **behavioral contrast** to describe the effects of (a) a schedule change that increases or decreases the rate of responding in one component of a multiple schedule of reinforcement, which is accompanied by (b) a change in the response rate in the opposite direction on another, unaltered component. That is, changing schedule requirements in one component may affect responding in other unaltered components (e.g., changing a mult FR 3/FI 30-sec to a mult FR 3/EXT may result in an increased rate of response in the FR component, the unaltered component).

For example, an analyst used the following intervention to decrease a student's disruptive behavior in the classroom and on the playground. She established an interval of time and delivered reinforcement at the end of the interval if the student had emitted no disruptive classroom behavior during the interval. With the occurrence of a disruptive behavior, the analyst immediately reset a timer, starting a new interval. The analyst introduced no intervention on playground behavior. This procedure effectively decreased the student's disruptive classroom behavior, and the number of playground disruptions remained low as well. With the effectiveness of differential reinforcement of other behavior (DRO; see Chapter 25), the analyst began to thin the classroom reinforcement schedule by extending the interval of time (e.g., 25-, 35-, and 45-min intervals). Schedule thinning effectively maintained the appropriate classroom behavior. However, the behavioral contrast of the increased occurrences of disruptive behavior on the playground surprised the analyst.

Although most research on behavioral contrast has been conducted with nonhuman subjects, Pierce and Epling (1995) concluded the phenomenon likely occurs in humans. We concur and advise applied behavior analysts to be aware of behavioral contrast effects in their practice.

Chained Schedules

A **chained schedule (chain)** is similar to a multiple schedule. The multiple and chained schedules have two or more basic schedule requirements that occur successively, and have a discriminative stimulus correlated with each independent schedule. A chained schedule differs from a multiple schedule in three ways. First, the basic schedules in a chain schedule always occur in a specific order, never in the random or unpredictable order of multiple schedules. Second, the behavior may be the same for all elements of the chain, or different behaviors may be required for different elements in the chain. Third, conditioned reinforcement for responding in the first element in a chain is the presentation of the second element; conditioned reinforcement for responding in the second element is presentation of the third element, and so on until all elements in the chain have been completed in a specific sequence. The last element normally produces unconditioned reinforcement in a laboratory setting, or unconditioned or conditioned reinforcement in applied settings.

Tarbox, Madried, Aguilar, Jacobo, and Schiff (2009) used two-component chain schedules (e.g., "mun" & "day" for "Monday"; "b" and "all" for "ball") to develop the echoic behavior of two children with autism and one child with developmental delay.

> During each chaining session, the therapist presented one echoic in three sequential trials as rapidly as feasible (i.e., the therapist presented a trial immediately after the completion of reinforcement for the previous trial). On the first trial, the therapist modeled the first component (e.g., "say 'mun'"). If the participant correctly imitated this component within 5 s, the therapist delivered reinforcement and immediately modeled the second component (e.g., "say 'day'"). If the participant imitated the second component correctly, the therapist delivered reinforcement, modeled the entire target echoic (e.g., "say 'Monday'"), and provided reinforcement for correct imitation. An incorrect response on any of the three trials resulted in a single repetition of that trial, after which the therapist resumed the sequence. (p. 903)

The following example shows an elaborate sequence of different behaviors that must occur in a specific order. To service a bicycle headset, the mechanic will complete a chain with 13 components: (1) Disconnect the front brake cable; (2) remove handlebar and stem; (3) remove front wheel; (4) remove locknut; (5) unscrew adjusting race; (6) take fork out of frame; (7) inspect races; (8) grease and replace bearing balls for lower stack; (9) grease and replace bearing balls for upper race; (10) grease threads of steering column; (11) put fork into frame and thread the screwed race; (12) return lock washer; (13) adjust and lock the headset. The final outcome (i.e., a clean, greased, and adjusted bicycle headset) is contingent on the completion of all 13 components. (For a detailed description of behavior chains and how applied behavior analysts help individuals learn new and more complex chains of behavior, see Chapter 23.)

Nondiscriminated Schedules of Reinforcement

Unlike with discriminated compound schedules of reinforcement, the presence of each component schedule in a nondiscriminated compound schedule of reinforcement is not associated with an S^D.

Mixed Schedules

A **mixed schedule of reinforcement (mix)** uses a procedure identical to that of the multiple schedules, except no discriminative stimuli signal the presence of the independent component schedules. For example, with a mix FR 10/FI 1 schedule, reinforcement sometimes occurs after the completion of 10 responses and sometimes occurs with the first response after a 1-minute interval from the preceding reinforcement.

The research of Tiger and Hanley (2005), presented in the section on multiple schedules, used a three-component mult schedule of reinforcement (SR+, EXT_1 and EXT_2) to teach preschool children how to recruit teacher attention at appropriate times. They also used a time-based rotation of those components without discriminative stimuli. This mix schedule (SR+, EXT_1, and EXT_2) "served as a baseline from which the influence of schedule-correlated stimuli on children's social approaches would be assessed." (p. 501)

Tandem Schedules

The **tandem schedule (tand)** operates like the chained schedule, except that discriminative stimuli are not used with the elements in the chain. In a tand FR 15/FI 2 schedule, after a participant makes 15 responses (FR 15), the first correct response following an elapse of 2 minutes (FI 2) produces reinforcement.

Jeffries et al. (2016) assessed the effectiveness of a CRF/FR 4 tandem schedule of reinforcement with children using an iPad platform. Essentially, a child examined a number (e.g., 7, 2) embedded within simulated eyes on the iPad with the direction to pick the matching number on a corresponding grid. If the number matched, a piece of coal appeared on the screen and a praise statement was emitted by the [iPad] narrator (CFR). After the child received four pieces of coal (FR 4), an interactive *Steam Train* game appeared, which the child was allowed to play for 1 min.

Although basic researchers have produced considerable data on the effects of mixed and tandem schedules, there is little applied research on these schedules. As applied behavior analysts continue to improve the science's knowledge base, how nondiscriminated compound schedules of reinforcement might contribute to assessment and intervention may become more apparent.

Schedules Combining Number of Responses and Time

Alternative Schedules

An **alternative schedule (alt)** provides an opportunity for reinforcement on two or more simultaneously available basic schedules of reinforcement. The first schedule completed provides reinforcement, regardless of which schedule component is met first. All schedule requirements start over following reinforcement. With an alt FR 50/FI 5-min schedule, reinforcement is delivered whenever either of these two conditions has been met: (a) 50 correct responses, provided the 5-minute interval of time has not elapsed; or (b) the first response after the elapse of 5 minutes, provided that fewer than 50 responses have been emitted.

For instance, a teacher using an alt FR 25/FI 3-min schedule of reinforcement assigns 25 math problems and assesses the student's correct and incorrect answers following the elapse of 3 minutes. If the student completes the 25 problems before the elapse of 3 minutes, the teacher checks the student's answers and provides a consequence consistent with the FR 25 schedule. However, if the ratio requirement of 25 math problems has not been completed after an elapse of 3 minutes, the first correct answer following the 3 minutes produces reinforcement. The alternative schedule offers the advantage of a second chance for reinforcement if the student has not met the FR requirement in a reasonable amount of time. The FI provides reinforcement for one response, and that one reinforced response might encourage continued responding with the new start of the FR requirement.

Conjunctive Schedules

A **conjunctive schedule (conj)** of reinforcement is in effect whenever reinforcement follows the completion of response requirements for two or more schedules of reinforcement. For example, a student behavior produces reinforcement when at least 2 minutes have elapsed and 50 responses have been made. This arrangement is a conj FI 2/FR 50 schedule of reinforcement. With the conjunctive schedule of reinforcement, the first response following the conclusion of the time interval produces reinforcement if the criterion number of responses has been completed.

A 14-year-old boy with autism had higher rates of aggression with two of his four therapists during instruction. The higher rates of aggression were directed toward the two therapists who previously worked with the boy at a different treatment facility. Progar et al. (2001) intervened to reduce the levels of aggression with the therapists from the different facility to the levels that occurred with the other two therapists in the current setting. The boy's aggression occurred in demand situations (e.g., making his bed) and was escape maintained. The initial intervention used three consequences: (1) a 10-minute chair time-out for attempts to choke, (2) escape extinction (see Chapter 24), and (3) differential reinforcement other behavior (DRO; see Chapter 25) for the omission of aggression during the 10-minute sessions. This intervention was identical to the treatment used with the boy at the other facility. It was ineffective in reducing the boy's aggression in the current setting.

Because of the ineffectiveness of the initial intervention, Progar and colleagues added a conj FR/VI-DRO schedule of reinforcement to their initial intervention. They delivered edible reinforcers contingent on completing a three-component task, such as dusting or straightening objects (i.e., an FR 3 schedule), and the omission of aggression for an average of every 2.5 minutes (i.e., a VI-DRO 150-sec). An occurrence of aggression reset the conj schedule. (*Note*: Resetting this conj schedule used a standard procedure because any occurrence of the problem behavior during a DRO interval immediately reset the time to the beginning of the interval.) Progar and colleagues demonstrated that the conj FR VI-DRO schedule produced a substantial reduction in aggression directed toward the two therapists previously from the other treatment facility.

Duvinsky and Poppen (1982) found that human performance on a conjunctive schedule is influenced by the ratio and interval requirements. When task requirements are high in relation to the interval requirements, people are likely to work steadily on the task throughout the time available. However, people are likely to engage in behaviors other than the task requirements when a large time interval and a low ratio requirement are present.

Table 13.1 describes and provides examples of the compound schedules of reinforcement presented above. Table 13.2 compares the defining dimensions of these schedules.

TABLE 13.1 Descriptions and Examples of Compound Schedules of Reinforcement

Discriminated Compound Schedules of Reinforcement

SCHEDULE	DESCRIPTION	EXAMPLE
Concurrent (conc)	Two or more schedules of reinforcement, each with a correlated S^D, operate independently and simultaneously for two or more behaviors.	A student is given two math worksheets, one blue with addition problems, one white with subtraction problems. The student can work on either or both types of problems, and earns a reinforcer for every 10 correct answers to addition problems and every 5 correct answers to subtraction problems. [conc FR 10/FR 5]
Multiple (mult)	Two or more basic schedules of reinforcement for the same behavior operate successively, usually in random sequence. An S^D is correlated with each component schedule. Reinforcement is delivered each time the response requirements are met for the schedule in effect.	During daily math practice activity, a student is given a blue worksheet or a white worksheet. Which color worksheet is assigned each day varies randomly. Regardless of color, each worksheet contains a mix of 100 addition problems. On blue worksheet days, the student earns a reinforcer for each 10 correct answers. On white worksheet days, the student receives a reinforcer each time he correctly answers a variable number (average, 5) of problems. [mult FR 10/VR 5]
Chained (chain)	Two or more basic schedules, each correlated with an S^D, operate in a specified sequence. The component schedules may require the same behavior or different behaviors. Meeting the response requirement in the first component presentation of the S^D for the second component must be accomplished, and so on until all elements in the chain have been completed. Completing the last element normally produces unconditioned or conditioned reinforcement in applied settings.	A student receives a worksheet containing 20 addition problems. When the student has correctly answered all of the problems, he receives a second worksheet, this one with 10 subtraction problems. After correctly answering the 10 subtraction problems, the student receives a reinforcer. [chain FR 20/FR 10]

Nondiscriminated Compound Schedules of Reinforcement

SCHEDULE	DESCRIPTION	EXAMPLE
Mixed (mix)	Two or more basic schedules are presented successively, usually in random sequence. Unlike with a multiple schedule, however, each component schedule is not signaled by an S^D.	A student is assigned daily math worksheets always of the same color. Sometimes the student receives a reinforcer after 10 correct answers; sometimes reinforcement follows the first correct answer after a 1-min interval has elapsed from the preceding reinforcement. [mix FR 10/FI 1]
Tandem (tand)	Two or more basic schedules operate in a specified sequence like a chained schedule. Unlike with a chained schedule, however, an S^D is not associated with each component schedule.	A student receives a worksheet with math problems. She earns a reinforcer for completing 15 problems, and then, the first problem completed after an elapse of 2 min also produces a reinforcer. [tand FR 15/FI 2]

Compound Schedules of Reinforcement Combining Number of Responses and Time

SCHEDULE	DESCRIPTION	EXAMPLE
Alternative (alt)	Reinforcement is obtained by meeting the response requirements of any of two or more simultaneously available component schedules. Reinforcement is delivered regardless of which schedule component is met. All component schedule requirements are reset following reinforcement.	A student receives a worksheet with math problems and is instructed to begin. If 15 correct answers are emitted prior to the elapse of 3 min, a reinforcer is delivered. If the FR requirement is not met after 3 min, the first correct answer from that time forward is reinforced. [FI 3-min/FR 15]
Conjunctive (conj)	Reinforcement follows the completion of response requirements for two or more simultaneously operating schedules of reinforcement.	A student receives a worksheet with math problems. After 15 problems have been answered correctly, the first correct answer after an elapse of 2 min produces reinforcement. [conj FR 15/FI 2]

TABLE 13.2 Defining Dimensions of Compound Schedules of Reinforcement

Dimension	Concurrent	Multiple	Chained	Mixed	Tandem	Alternative	Conjunctive
Number of component schedules	2 or more	2 or more	2 or more	2 or more	2 or more	2 or more	2 or more
Number of response classes involved	2 or more	1	1 or more	1	1 or more	1	1
Discriminative stimuli or cues for each component schedule	Possible	Yes	Yes	No	No	Possible	Possible
Successive presentation of component schedules	No	Yes	Yes	Yes	Yes	No	No
Simultaneous presentation of component schedules	Yes	No	No	No	No	Yes	Yes
Reinforcement limited to final component schedule	No	No	Yes	No	Yes	No	Yes
Must meet response requirements of each component schedule for reinforcement	Yes	Yes	No	Yes	No	Yes	No

PERSPECTIVES ON USING SCHEDULES OF REINFORCEMENT IN APPLIED SETTINGS

Applied Research with Intermittent Schedules

Basic researchers have systematically analyzed the effects of intermittent schedules of reinforcement on behavior (e.g., Ferster & Skinner, 1957). Their results have produced well-established schedule effects. These schedule effects have strong generality across many species, response classes, and laboratories. However, a review of the applied literature on schedule effects (e.g., *Journal of Applied Behavior Analysis*, 1968 to 2016) will show that applied behavior analysts have not embraced the analysis of schedule effects with enthusiasm, as have basic researchers. Consequently, schedule effects have not been documented clearly in applied settings. Uncontrolled variables in applied settings, such as the following, influence a participant's sensitivity and insensitivity to the schedule of reinforcement:

1. Instructions given by the applied behavior analyst, self-instructions, and environmental aids (e.g., calendars, clocks) make human participants resistant to temporal schedule control.
2. Past histories of responding to intermittent schedules of reinforcement can affect current schedule sensitivity or insensitivity.
3. Immediate histories from schedules of reinforcement may affect current schedule performances more than remote past histories.
4. Sequential responses required in many applied applications of intermittent schedules of reinforcement (e.g., work leading to the paycheck, studying for a pop quiz) are uncommon applications of schedules of reinforcement, particularly with interval schedules.
5. Uncontrolled establishing operations in conjunction with schedules of reinforcement in applied settings will confound schedule effects.

Some well-established schedule effects found in basic research were presented earlier in this chapter. Applied behavior analysts, however, should use caution in extrapolating these effects to applied settings, for the following reasons:

1. Most applied applications of schedules of reinforcement only approximate true laboratory schedules of reinforcement, especially the interval schedules that may occur rarely in natural environments (Nevin, 1998).
2. Many uncontrolled variables in applied settings will influence a participant's sensitivity and insensitivity to the schedule of reinforcement (Madden, Chase, & Joyce, 1998).

Applied Research with Compound Schedules

Although a considerable body of applied research on multiple, concurrent, and chained schedules has been reported, scant research exists on the effects of nondiscriminated compound schedules of reinforcement in applied settings. We encourage applied behavior analysts to explore the use of mixed, tandem, alternative, and conjunctive schedules. Doing so will reveal the effects of these compound schedules in applied settings and their interactions with other environmental variables (e.g., antecedent stimuli, motivating operations) (Lattal & Neef, 1996).

Applied Research with Adjunctive Behavior

This chapter has stressed the effects of schedules of reinforcement on the specific behaviors that produce reinforcement. Other behaviors can occur when an individual responds to a given contingency of reinforcement. These other behaviors occur independently of schedule control. Typical examples of such behaviors include normal time fillers, such as doodling, smoking, idle talking, and

drinking. Such behaviors are called **adjunctive behaviors**, or schedule-induced behaviors, when the frequency of these time-filling behaviors increases as a side effect of other behaviors maintained by a schedule of reinforcement (Falk, 1961, 1971).

A substantial body of experimental literature is available on many types of adjunctive behaviors with nonhuman subjects (see reviews: Staddon, 1977; Wetherington, 1982) and some basic research with human subjects (e.g., Kachanoff, Leveille, McLelland, & Wayner 1973; Lasiter, 1979). Common diverse examples of adjunctive behaviors observed in laboratory experiments include aggression, defecation, pica, and wheel running. Some common excessive human problem behaviors might develop as adjunctive behaviors (e.g., the use of drugs, tobacco, caffeine, and alcohol; overeating; nail biting; self-stimulation; and self-abuse). These potentially excessive adjunctive behaviors are socially significant, but the possibility that such excesses are developed and maintained as adjunctive behaviors has been essentially ignored in applied behavior analysis.

Foster (1978), in an extended communication to the readership of the *Journal of Applied Behavior Analysis*, reported that applied behavior analysts have neglected the potentially important area of adjunctive behavior. He stated that applied behavior analysis does not have a data or knowledge base for adjunctive phenomena. Similarly, Epling and Pierce (1983) called for applied behavior analysts to extend the laboratory-based findings in adjunctive behavior to the understanding and control of socially significant human behavior. To our knowledge, Lerman, Iwata, Zarcone, and Ringdahl's (1994) article provides the only research on adjunctive behavior published in the *Journal of Applied Behavior Analysis* from 1968 through 2016. Lerman and colleagues provided an assessment of stereotypic and self-injurious behavior as adjunctive responses. Data from this preliminary study suggest that intermittent reinforcement did not induce self-injury, but with some individuals, stereotypic behavior showed characteristics of adjunctive behavior.

Foster (1978) and Epling and Pierce (1983) cautioned that many teachers and therapists may apply interventions directly to adjunctive behaviors rather than to the variables functionally related to their occurrence. These direct interventions may be futile and costly in terms of money, time, and effort because adjunctive behaviors appear resistant to interventions using operant contingencies.

The condition under which adjunctive behaviors are developed and maintained is a major area for future research in applied behavior analysis. Applied research directed to adjunctive behaviors will advance the science of applied behavior analysis and will provide an important foundation for improved practices in therapy and instruction.

SUMMARY

Intermittent Reinforcement

1. A schedule of reinforcement is a rule that establishes the probability that a specific occurrence of a behavior will produce reinforcement.
2. Continuous reinforcement and extinction provide the boundaries for all other schedules of reinforcement.
3. Only selected occurrences of behavior produce reinforcement with an intermittent schedule of reinforcement.
4. Continuous reinforcement (CRF) is used to strengthen behavior, primarily during the initial stages of learning new behaviors.
5. Intermittent reinforcement (INT) is used to maintain an established behavior, especially during maintenance stages of learning.

Defining Basic Intermittent Schedules of Reinforcement

6. Ratio schedules require a number of responses before one response produces reinforcement.
7. Interval schedules require an elapse of time before a response produces reinforcement.

Schedule Effects and Consistency of Performance

8. A fixed ratio (FR) schedule requires the completion of a specified number of responses before a response produces reinforcement.
9. FR schedules produce a typical pattern of responding: (a) After the first response of the ratio requirement, the participant completes the required responses with little hesitation between responses; and (b) a postreinforcement pause follows reinforcement.
10. The size of the ratio and the reinforcement magnitude can influence the duration of the postreinforcement pause.
11. A variable ratio (VR) schedule of reinforcement requires the completion of a variable number of responses before a response produces reinforcement. A number representing the average (e.g., mean) number of responses required for reinforcement identifies the VR schedule.
12. VR schedules produce consistent, steady rates of response. They typically do not produce a postreinforcement pause, as do FR schedules.
13. A fixed interval (FI) schedule provides reinforcement for the first response following the elapse of a specific, constant duration of time since the last reinforced response. The reinforcer is available after the fixed time interval has elapsed, and it remains available until the first response is emitted.
14. FI schedules typically produce a postreinforcement pause in responding during the early part of the interval. An initially slow, but accelerating, rate of response is evident toward the end of the interval, usually reaching a maximum rate just before delivery of the reinforcer. This gradually accelerating rate of response toward the

end of the interval is called an FI scallop because of the rounded curves that are shown on a cumulative graph.

15. A variable interval (VI) schedule provides reinforcement for the first response following the elapse of a variable duration of time.

16. A VI schedule of reinforcement tends to produce a low to moderate stable rate of response. On a cumulative graph, the slope of the VI schedule appears uniform with few pauses in responding.

17. When a limited hold is added to an interval schedule, reinforcement remains available for a finite time following the elapse of the FI or VI interval.

Thinning Intermittent Reinforcement

18. Two procedures are used commonly for schedule thinning. First, the response ratio or the duration of the time interval is gradually increased. Instructions can be used to clearly communicate the schedule of reinforcement, facilitating a smooth transition during the thinning process. Instructions include rules, directions, and signs.

19. To adjust the thinning process, and to avoid the loss of previous improvements, use small increments of schedule changes.

20. Ratio strain can result from abrupt increases in ratio requirements when moving from denser to thinner reinforcement schedules.

Variations on Basic Intermittent Schedules of Reinforcement

21. Differential reinforcement of high rates (DRH) and differential reinforcement of low rates (DRL) are variations of ratio schedules that make reinforcement contingent on responses occurring above or below criterion response rates.
    - DRH schedules produce higher rates of responding.
    - DRL schedules produce lower rates of responding.

22. The differential reinforcement of diminishing rates schedule provides reinforcement at the end of a predetermined time interval when the number of responses is less than a criterion that is gradually decreased across time intervals based on the individual's performance.

23. A lag schedule of reinforcement is one method to increase response variability. Reinforcement on a lag schedule is contingent on a response differing in some predetermined way (e.g., different topography, different sequence) from one or more responses that preceded it.

24. A progressive ratio schedule of reinforcement is a variation of the fixed ratio (FR) schedule in which the ratio requirements increase incrementally within the session.

Compound Schedules of Reinforcement

25. Continuous reinforcement, the four basic intermittent schedules of reinforcement, differential reinforcement of rates of responding, and extinction, when combined, produce compound schedules of reinforcement.

26. The basic component schedules within a compound schedule (a) can occur successively or simultaneously; (b) can occur with or without discriminative stimuli; and (c) make reinforcement contingent on meeting the response requirements of each component schedule independently, or in combination with compound schedules.

27. A multiple schedule (mult) presents two or more basic schedules of reinforcement in an alternating, usually random, sequence. The basic schedules within the multiple schedule occur successively and independently. A discriminative stimulus is correlated with each basic schedule, and that stimulus is present as long as the schedule is in effect.

28. Behavioral contrast describes the effects of (a) a schedule change that increases or decreases the rate of responding in one component of a multiple schedule of reinforcement, which is accompanied by (b) a change in the response rate in the opposite direction on another unaltered component.

29. A concurrent schedule (conc) of reinforcement occurs when (a) two or more contingencies of reinforcement (b) operate independently and simultaneously (c) for two or more behaviors. Each schedule is correlated with a discriminative stimulus.

30. A chained schedule (chain) is similar to a multiple schedule. The multiple and chained schedules have two or more basic schedule requirements that occur successively, and have a discriminative stimulus correlated with each independent schedule.

31. A mixed schedule of reinforcement (mix) uses a procedure identical to that of the multiple schedules, except no discriminative stimuli signal the presence of the independent component schedules.

32. A tandem schedule (tand) operates like the chained schedule, except that discriminative stimuli are not used with the elements in the chain.

33. An alternative schedule (alt) provides an opportunity for reinforcement on two or more simultaneously available basic schedules of reinforcement. The first schedule completed provides reinforcement, regardless of which schedule component is met first. All schedule requirements start over following reinforcement.

34. A conjunctive schedule (conj) of reinforcement is in effect whenever reinforcement follows the completion of response requirements for two or more schedules of reinforcement.

Perspectives on Using Schedules of Reinforcement in Applied Settings

35. Well-established schedule effects found in basic research were presented in this chapter. Applied behavior analysts, however, should use caution in extrapolating these effects to applied settings.

36. Applied behavior analysts should include an analysis of the basic intermittent schedules and the compound schedules in their research agendas. A better understanding of the schedule effects in applied settings will advance the development of applied behavior analysis and its applications.

37. The conditions under which adjunctive behaviors (e.g., time fillers, such as doodling, smoking, idle talking, and drinking) are developed and maintained are important areas for future research in applied behavior analysis.

KEY TERMS

adjunctive behaviors
alternative schedule (alt)
behavioral contrast
chained schedule (chain)
compound schedule of reinforcement
concurrent schedule (conc)
conjunctive schedule (conj)
continuous reinforcement (CRF)
differential reinforcement of diminishing rates (DRD)
differential reinforcement of high rates (DRH)
differential reinforcement of low rates (DRL)
fixed interval (FI)
fixed ratio (FR)
intermittent schedule of reinforcement (INT)
lag schedule
limited hold
matching law
mixed schedule of reinforcement (mix)
multiple schedule (mult)
postreinforcement pause
progressive schedule of reinforcement
progressive-ratio (PR) schedule of reinforcement
ratio strain
schedule of reinforcement
schedule thinning
tandem schedule (tand)
variable interval (VI)
variable ratio (VR)

MULTIPLE-CHOICE QUESTIONS

1. Every time Baxter raised his hand in class, Mrs. Tulley called on him. What schedule of reinforcement is Baxter's hand raising on?

a. Intermittent reinforcement

b. Fixed-interval

c. Continuous reinforcement

d. Fixed-ratio 2

Hint: (See "Continuous Reinforcement")

2. Timmy has just recently begun clearing his plate after he finishes his meal. Timmy's mother really likes this new behavior and would like to see more of it. Which schedule of reinforcement should Timmy's mother use to strengthen Timmy's new behavior of clearing his plate?

a. Variable-ratio 8

b. Intermittent reinforcement

c. Fixed-ratio 10

d. Continuous reinforcement

Hint: (See "Intermittent Reinforcement")

3. Which schedule of reinforcement occurs when two or more contingencies of reinforcement operate independently and simultaneously for two or more behaviors?

a. Variable-ratio schedules of reinforcement

b. Concurrent schedules of reinforcement

c. Fixed-interval schedules of reinforcement

d. Multiple schedules of reinforcement

Hint: (See "Compound Schedules of Reinforcement")

4. Which schedule of reinforcement presents two or more basic schedules of reinforcement in an alternating (usually random) sequence usually occurring successively and independently with no correlated discriminative stimuli?

a. Multiple schedules of reinforcement

b. Alternative schedules of reinforcement

c. Variable-interval schedules of reinforcement

d. Mixed schedules of reinforcement

Hint: (See "Discriminated Compound Schedules of Reinforcement")

5. Which schedule of reinforcement produces a post-reinforcement pause?
   a. Fixed ratio
   b. Continuous reinforcement schedule
   c. Tandem schedules of reinforcement
   d. Variable-interval schedules of reinforcement

   Hint: (See "Defining Basic Intermittent Schedules of Reinforcement")

6. To gradually reduce a high-frequency behavior, which schedule of reinforcement is the most appropriate?
   a. Differential reinforcement of high rates of responding
   b. Differential reinforcement of low rates of responding
   c. Fixed-ratio schedules of reinforcement
   d. Concurrent schedules of reinforcement

   Hint: (See "Schedules of Differential Reinforcement of Rates of Responding")

7. Schedule thinning refers to:
   a. Gradually decreasing the response ratio or the duration of the time interval
   b. Quickly increasing the response ratio or the duration of the time interval
   c. Gradually increasing the response ratio or the duration of the time interval
   d. Moving gradually from an intermittent schedule to a continuous reinforcement schedule.

   Hint: (See "Thinning Intermittent Reinforcement")

8. Common behavioral characteristics associated with ________ include avoidance, aggression, and unpredictable pauses in responding.
   a. Ratio strain
   b. Post-reinforcement pause
   c. Scallop effect
   d. Fixed interval

   Hint: (See "Thinning Intermittent Reinforcement")

9. Which schedule of reinforcement is identical to the chained schedule except it does not use discriminative stimuli with the elements in the chain?
   a. Mixed schedule of reinforcement
   b. Fixed-ratio schedule of reinforcement
   c. Tandem schedule of reinforcement
   d. Chained schedule of reinforcement

   Hint: (See "Discriminated Compound Schedules of Reinforcement"; "Nondiscriminated Schedules of Reinforcement")

10. When the frequency of a time-filling behavior increases as a side effect of some other behavior maintained by a schedule of reinforcement, this is referred to as:
    a. Ratio strain
    b. Post-reinforcement pause
    c. Schedule induced
    d. Schedule thinning

    Hint: (See "Applied Research with Adjunctive Behavior")

ESSAY-TYPE QUESTIONS

1. Discuss the term schedules of reinforcement. In your discussion, be sure to describe what it is, and discuss the "boundaries for all other schedules of reinforcement" identified in the chapter.

   Hint: (See "Defining Basic Intermittent Schedules of Reinforcement")

2. Explain in detail how you would utilize schedules of reinforcement for teaching a student a new skill and, once acquired, what schedule you would use to maintain it.

   Hint: (See "Defining Basic Intermittent Schedules of Reinforcement")

3. Compare and contrast the four basic intermittent schedules of reinforcement.

   Hint: (See "Defining Basic Intermittent Schedules of Reinforcement")

4. Identify and describe the schedule of effects, specifically the issues of consistency of performance and rate of response, for each of the four basic intermittent schedules of reinforcement.

   Hint: (See "Defining Basic Intermittent Schedules of Reinforcement")

5. Describe the process of how you would use a ratio schedule to (1) teach a student a new skill and (2) move the skill into maintenance.

   Hint: (See "Thinning Intermittent Reinforcement")

6. During the treatment condition, a student became avoidant of work, began engaging in physical aggression and random pauses in responding. Upon inquiry, you discover that the staff was in the process of thinning the schedule of reinforcement the student was on. Identify the term that describes these phenomena and discuss how it occurred. How can it be remedied?

   Hint: (See "Thinning Intermittent Reinforcement")

7. Julia is a 9-year-old student at Blackmore Elementary. You have been asked by your professor to help her teacher reduce the frequency of Julia's hand raising in class using a differential reinforcement strategy. The teacher has recorded Julia raising her hand an average of 57 times an hour. Discuss which differential reinforcement procedure you would use, why you would use it and how it would work.

   Hint: (See "Schedules of Differential Reinforcement of Rates of Responding")

8. Identify and discuss the discriminated schedules of reinforcement.

Hint: (See "Discriminated Compound Schedules of Reinforcement")

9. Identify and discuss the nondiscriminated schedules of reinforcement.

Hint: (See "Nondiscriminated Schedules of Reinforcement")

10. Discuss and explain two different ways applied behavior analysts use progressive schedules of reinforcement.

Hint: (See "Progressive Schedules of Reinforcement")

NOTES

1. While we define the PR schedule of reinforcement as increasing response requirements within the session, as introduced by Hodos (1961), we recognize that applied behavior analysts may increase response requirements over successive sessions rather than within the session (e.g., DeLeon, Iwata, Goh, & Worsdell, 1997; Wilson & Gratz, 2016). Jarmolowizc and Lattal (2010) suggested the label progressive fixed-ratio (PFR) for schedules in which response requirements are increased from one session to the next.
2. An escalating schedule of reinforcement, in which consecutive responses produce an ever-increasing magnitude of reinforcement (e.g., Romanowich & Lamb, 2010, 2015), should not be confused with a PR schedule of reinforcement.

PART 5

Punishment

Punishment is a basic principle of learning. Although often considered bad—the unfortunate evil twin to reinforcement—punishment has as much importance to learning as reinforcement does. Learning from consequences that produce pain, discomfort, or the loss of reinforcers has survival value for the individual and for the species.

As with reinforcement, stimulus changes that function as consequences in punishment contingencies can be described by two types of operations: A new stimulus is presented or an existing stimulus is removed. Chapter 14 begins by defining the basic principle of punishment and distinguishes positive punishment and negative punishment based on the operation of the response-suppressing consequence. The remainder of the chapter focuses on the side effects and limitations of punishment, factors that influence the effectiveness of punishment, examples of interventions involving positive punishment, guidelines for using punishment effectively, and ethical considerations related to using punishment in applied settings. Chapter 15 describes two behavior change tactics based on negative punishment: time-out from positive reinforcement and response cost. Examples show how time-out and response cost reduce or eliminate undesired behavior, and guidelines provide direction for effectively implementing these tactics.

CHAPTER 14

Positive Punishment

LEARNING OBJECTIVES

- Define and discuss the definitions of positive and negative punishment.
- Define and discuss factors that influence the effectiveness of punishment.
- Identify and explain the potential side effects of punishment use.
- List examples of positive punishment and describe their respective procedures.
- Discuss the guidelines for the use of punishment as an intervention.
- Identify and discuss ethical considerations regarding the use of punishment.
- Explain the current state of knowledge regarding the use of punishment as a treatment intervention.

Have you ever stubbed your toe while walking too fast in a darkened room and then walked slowly the rest of the way to the light switch? Have you ever smashed your finger with a hammer and then refrained from holding the nail in quite the same way before swinging the hammer again? Have you ever left a sandwich unattended at a beach party, watched a seagull fly away with it, and then did not turn your back on the next treat you pulled from the picnic basket? If you have had these or similar experiences, you have been the beneficiary of punishment.

It may seem strange to refer to people who stubbed their toe, smashed a finger, or lost a tasty sandwich as benefiting from the experience, as opposed to referring to that person as a "victim" of punishment. Although many people consider punishment a bad thing—reinforcement's evil counterpart—punishment is as important to learning as reinforcement. Learning from consequences that produce pain, discomfort, or the loss of reinforcers has survival value for the individual organism and for our species. Punishment teaches us not to repeat behaviors that cause harm. Fortunately, it usually does not take too many stubbed toes or smashed fingers to reduce the occurrence of the behaviors that produced those outcomes.

Although punishment is a natural phenomenon and one of the basic principles of operant conditioning, it is poorly understood and frequently misapplied, and its application is controversial. At least some of the misunderstandings and controversy surrounding the use of punishment in behavior change programs can be attributed to practitioners who, on the one hand, confuse punishment as an empirically derived principle of behavior with its commonplace usage and its legal connotations; and, on the other hand, misapply punishment altogether.

Punishment is generally understood to be the application of aversive consequences—such as physical pain, psychological hurt, and the loss of privileges, or fines—for the purpose of teaching a lesson to a person who has misbehaved so she will not repeat the misdeed. Punishment is sometimes used as an act of retribution on the part of the person or agency administering the punishment, or to provide "a lesson for others" on how to behave. Punishments meted out by the legal system, such as jail time and fines, are often considered a process by which convicted lawbreakers must repay their debt to society.

These everyday and legal notions of punishment, although having various degrees of validity and utility within their own contexts, have little, if anything, to do with punishment as a principle of behavior. In the everyday connotation of punishment, most people would agree that a teacher who sends a student to the principal's office for misbehaving in class, or a police officer who issues a ticket to a speeding motorist, has punished the offender. However, as a principle of behavior, punishment is not about *punishing the person*; punishment is a *response → consequence contingency* that suppresses the future frequency of similar responses. From the perspective of both the science and the practice of behavior analysis, the trip to the principal's office did not punish fooling around in class unless the future rate at which the student fools around in class decreases as a function of his trip to the principal's office, and the police officer's ticket did not punish speeding unless the motorist drives above the speed limit less often than she did before receiving the ticket.

In this chapter, we define punishment, identify factors that influence the effectiveness of punishment, discuss its side effects and limitations, describe examples of several behavior change tactics that incorporate punishment, present guidelines for using punishment effectively, and discuss ethical considerations in the use of punishment. In the chapter's concluding section, we remind behavior analysts of punishment's natural and necessary role in learning and underscore the need for more basic and applied research on punishment, and state that a wide variety of differential reinforcement options, stimulus control procedures, time-out from positive reinforcement, response blocking, and contingent restraint, be employed in lieu of treatments featuring unconditioned punishers as default technologies.

DEFINITION AND CHARACTERISTICS OF PUNISHMENT

This section presents the functional relation that defines punishment, describes the two operations by which punishment can be implemented, identifies the types of consequences that function as punishers, explains the discriminative effect of punishment, and describes the recovery from the punishment effect.

Operation and Defining Effect of Punishment

Like reinforcement, punishment is a two-term functional relation defined by its effects on the frequency of behavior. **Punishment** has occurred when a response is followed immediately by a stimulus change that decreases the future frequency of that type of behavior.

An early study by Hall and colleagues (1971) provides a straightforward example of punishment. Andrea, a 7-year-old girl with hearing impairments, "pinched and bit herself, her peers, the teacher, and visitors to the classroom at every opportunity and was so disruptive the teacher reported that academic instruction was impossible" (p. 24). During an initial 6-day baseline period, Andrea bit or pinched an average of 71.8 times per day. Whenever Andrea bit or pinched anyone during the intervention condition, her teacher immediately pointed at her with an outstretched arm and shouted "No!" On the first day of intervention, the frequency of Andrea's biting and pinching decreased substantially (see Figure 14.1). Her aggressive behavior followed a downward trend during the initial intervention phase, ending in an average of 5.4 incidents per day. A 3-day return to baseline conditions resulted in Andrea's biting and pinching at a mean rate of 30 times per day. When the teacher reinstated the intervention, pointing her finger and stating "No!" each time Andrea pinched or bit, Andrea's problem behavior dropped to 3.1 incidents per day. During the second intervention phase, the teacher reported that Andrea's classmates were no longer avoiding her, perhaps because their behavior of being near Andrea was punished less often by bites and pinches.

It is important to point out that punishment is defined neither by the actions of the person delivering the consequences (in the case of socially mediated punishment) nor by the nature of those consequences.[1] A decrease in the future occurrence of the behavior must be observed before a consequence-based intervention qualifies as punishment. The intervention that proved successful in reducing the frequency of Andrea's biting and pinching—her teacher's pointed finger and "No!"—is classified as a punishment-based treatment only because of its suppressive effects. If Andrea had continued to bite and pinch at the baseline level of responding when the intervention was applied, her teacher's pointing and "No!" would not have been punishment.

Because the presentation of punishers often evokes behavior incompatible with the behavior being punished, the immediate suppressive effects of punishment can easily be overestimated. Michael (2004) explained and provided a good example:

> The decreased frequency of the punished response that is due to its having been followed by punishment will not be seen until after behavior evoked by the punishing stimulus changes has ceased. Because the evocative effect of the punishment stimulus change (as a respondent unconditioned or conditioned elicitor or as operant S^D or MO) is in the same direction as the future change due to punishment as a response consequence (weakening of the punished behavior), the former can be easily misinterpreted as the latter. For example, when a small child's misbehavior is followed by a severe reprimand, the misbehavior will cease immediately, but primarily because the reprimand controls behavior incompatible with the misbehavior—attending to the adult who is doing the reprimanding, denying responsibility for the misbehavior, emotional behavior such as crying, etc. This sudden and total cessation of the misbehaver does not imply, however, that the future frequency of its occurrence has been reduced, which would be the true effect of punishment. (pp. 36–37)

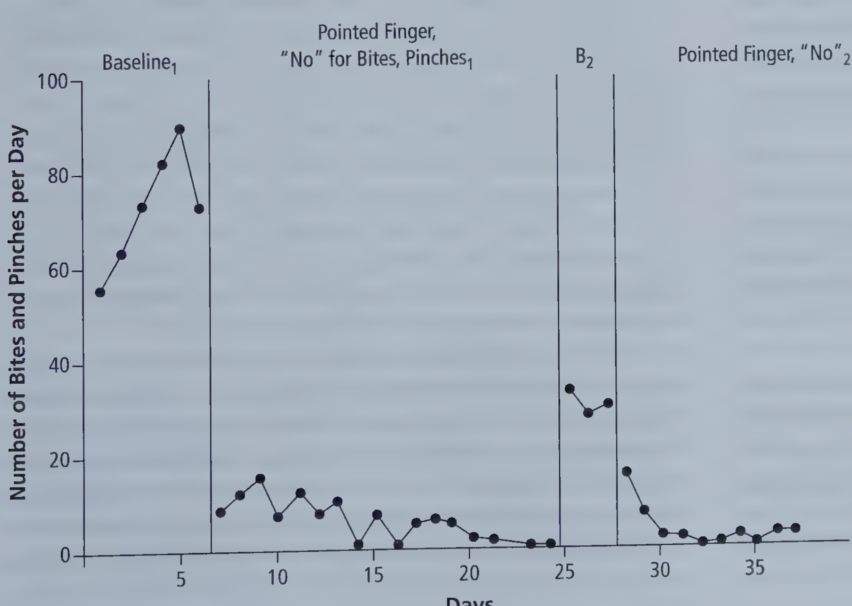


Figure 14.1 Number of bites and pinches by a 7-year-old girl during baseline and punishment ("No" plus pointing) conditions.

From "The Effective Use of Punishment to Modify Behavior in the Classroom," by R. V. Hall, S. Axelrod, M. Foundopoulos, J. Shellman, R. A. Campbell, and S. S. Cranston, 1971, *Educational Technology, 11*(4), p. 25. Copyright 1971 by Educational Technology. Used by permission.

Extinction effects caused by withholding reinforcement for the problem behavior (a procedure that should be included in punishment-based interventions whenever possible) contribute to the difficulty of determining the effectiveness of punishment—or even whether the punishment contingency is responsible for reductions in responding. However, rapid reductions in responding characteristic of punishment can usually be distinguished from gradual reductions in responding typical of extinction. Compare the immediate decrease in Andrea's biting and pinching from baseline to punishment condition in Figure 14.1 with the gradual decrease in responding typical of extinction shown in Figure 24.4.

Positive Punishment and Negative Punishment

Like reinforcement, punishment can be accomplished by either of two types of stimulus change operations. **Positive punishment** occurs when the presentation of a stimulus (or an increase in the intensity of an already present stimulus) immediately following a behavior results in a decrease in the frequency of the behavior. Stubbing one's toe on a chair leg is an example of positive punishment—if it suppresses the frequency of the behavior that preceded the toe stub—because the painful stimulation is best described as the presentation of a stimulus. Behavior change tactics based on positive punishment involve the contingent presentation of a stimulus immediately following occurrences of the target behavior. Andrea's teacher constituted positive punishment; the teacher's pointed finger and "No!" were stimuli presented or added to Andrea's environment.

Negative punishment has occurred when the removal of an already present stimulus (or a decrease in the intensity of an already present stimulus) immediately following a behavior results in a decrease in the future frequency of the behavior. The beach party attendee's behavior of turning his back on his food was negatively punished when the seagull flew off with his sandwich. For a stimulus change to function as negative punishment, which amounts to the removal of a positive reinforcer, a "motivating operation for the reinforcer must be in effect, otherwise removing it will not constitute punishment" (Michael, 2004, p. 36). A seagull's flying off with a hungry person's sandwich would function as punishment for his inattentiveness, but perhaps have little effect on the behavior of a person who has eaten his fill and set the sandwich down.

Behavior change tactics based on negative punishment involve the contingent loss of available reinforcers immediately following a behavior (i.e., response cost, a procedure akin to a fine) or the removal of the opportunity to acquire additional reinforcers for a period of time (i.e., time-out from reinforcement, a procedure akin to being sidelined during a game). Negative punishment is the focus of Chapter 15.

As with reinforcement, the modifiers *positive* and *negative* when used with *punishment* connote neither the intention nor the desirability of the behavior change produced; they specify only that the stimulus change that functioned as the punishing consequence is best described as the presentation of a new stimulus (positive punishment) or the termination (or reduction in intensity or amount) of an already present stimulus (negative punishment). As Michael (2004) noted, positive punishment and negative punishment are susceptible to misunderstanding.

> Assuming that you must receive either positive or negative punishment, which would you prefer? As with reinforcement, you should certainly not decide until you know specifically what each consists of. Would you prefer negative reinforcement or positive punishment? Of course, negative reinforcement. (p. 37)

Positive punishment and negative reinforcement are frequently confused as well. Because aversive events are associated with positive punishment and with negative reinforcement, the umbrella term *aversive control* is often used to describe interventions involving either or both of these two principles. Distinguishing between the two principles is difficult when the same aversive event is involved in concurrent positive punishment and negative reinforcement contingencies. For example, Baum (2017) described how the application and threat of physical beatings might condition the behavior of people living in a police state.

> If speaking out results in a beating, then speaking out is positively punished. If lying avoids a beating, then lying is negatively reinforced. The two tend to go hand-in-hand, because if one activity is punished, some alternative action will avoid punishment. (p. 163)

The keys to identifying and distinguishing concurrent positive punishment and negative reinforcement contingencies involving the same aversive stimulus event are (a) recognizing the opposite effects the two contingencies have on the future frequency of behavior and (b) realizing that two different behaviors must be involved because the same consequence (i.e., stimulus change) cannot serve as positive punishment and negative reinforcement for the same behavior. In a positive punishment contingency, the stimulus is absent prior to a response and is presented (or increased in intensity) as a consequence; in a negative reinforcement contingency, the stimulus is present prior to a response and is removed (or decreased in intensity) as a consequence. For example, positive punishment and negative reinforcement contingencies operated concurrently in a study by Azrin, Rubin, O'Brien, Ayllon, and Roll (1968) in which adults wore an apparatus throughout the normal working day that automatically produced a 55-dB tone contingent on slouching (sustained rounding of the shoulders or upper back for 3 seconds) and immediately terminated the tone when they straightened their shoulders. Slouching produced a tone (positive punishment), straightening the shoulders escaped the tone (negative reinforcement), and nonslouching avoided a tone (negative reinforcement).

Threatening a person with punishment if she engages in a behavior should not be confused with punishment. Punishment is a behavior → consequence relation, and a threat of what will happen if a person behaves in a certain way is an antecedent event to the behavior. When the threat of punishment suppresses behavior, it may be due to the threat functioning as an establishing operation that evokes alternative behaviors that avoid the threatened punishment.

Types of Punishers

A **punisher** is a stimulus change that immediately follows a response and decreases future occurrences of that type of behavior. The stimulus change in a positive punishment contingency may be called a *positive punisher* or, more simply, a *punisher*. Likewise the term *negative punisher* can be applied to a stimulus change involved in negative punishment. Though technically accurate, the term *negative punisher* is clumsy because it refers to a positive reinforcer that is removed contingent on occurrences of the target behavior. Therefore, when most behavior analysts use the term *punisher*, they are referring to a stimulus whose presentation functions as punishment (i.e., a positive punisher). As with reinforcers, punishers can be classified as unconditioned or conditioned.

Unconditioned Punishers

An **unconditioned punisher** (also called *primary punisher* or *unlearned punisher*) is a stimulus whose presentation functions as punishment without having been paired with any other punishers. Because unconditioned punishers are the product of the evolutionary history (phylogeny) of a given species, all biologically intact members of that species are susceptible to punishment by the same unconditioned punishers. Painful stimulation such as that caused by physical trauma to the body, certain odors and tastes, physical restraint, loss of bodily support, and extreme muscular effort are examples of stimulus changes that typically function as unconditioned punishers for humans (Michael, 2004). Like unconditioned reinforcers, unconditioned punishers are "phylogenetically important events that bear directly on fitness" of the organism (Baum, 2017, p. 67).

However, virtually any stimulus to which an organism's receptors are sensitive—light, sound, and temperature, to name a few—can be intensified to the point that its delivery will suppress behavior even though the stimulus is below levels that cause tissue damage (Bijou & Baer, 1965). Unlike unconditioned reinforcers, such as food and water, whose effectiveness depends on a relevant establishing operation, under most conditions many unconditioned punishers will suppress any behavior that precedes their onset. An organism does not have to be "deprived of electric stimulation" for the onset of electric shock to function as punishment. (However, the behavior of an organism that has just received many shocks at a specified voltage level in a short period, particularly shocks of mild intensity, may be relatively unaffected by another shock at a similar intensity.)

Conditioned Punishers

A **conditioned punisher** (also called *secondary punisher* or *learned punisher*) is a stimulus change whose presentation functions as punishment as a result of a person's conditioning history. A conditioned punisher acquires the capability to function as a punisher through stimulus–stimulus pairing with one or more unconditioned, or previously conditioned, punishers. For example, as a result of its onset at or very near the same time as an electric shock, a previously neutral stimulus change, such as an audible tone, will become a conditioned punisher capable of suppressing behavior that immediately precedes the tone when it occurs later without the shock (Hake & Azrin, 1965).[2] If a conditioned punisher is repeatedly presented without the punisher(s) with which it was initially paired, its effectiveness as punishment will wane until it is no longer a punisher.

Previously neutral stimuli can also become conditioned punishers for humans without direct physical pairing with another punisher through a pairing process Alessi (1992) called *verbal analog conditioning*. This is similar to the example of verbal pairing described in Chapter 11 for the conditioning of a conditioned reinforcer in which Engelmann (1975) showed cut-up pieces of yellow construction paper to a group of preschool children and told them, "These pieces of yellow paper are what big kids work for" (pp. 98–100). From that point on, many children began working extra hard for yellow pieces of paper. Miltenberger (2016) gave the example of a carpenter telling his apprentice that if the electric saw starts to make smoke, the saw motor may become damaged or the blade may break. The carpenter's statement establishes smoke from the saw as a conditioned punisher capable of decreasing the frequency of any behaviors that immediately precede the smoke (e.g., pushing too forcefully on the saw, holding the saw at an inappropriate angle).

A stimulus change that has been paired with numerous forms of unconditioned and conditioned punishers becomes a **generalized conditioned punisher**. Reprimands ("No!" "Don't do that!") and disapproving gestures (e.g., scowls, head shakes, frowns) are generalized conditioned punishers for many people because they have been paired repeatedly with a wide range of unconditioned or conditioned punishers (e.g., burned finger, loss of privileges). As with generalized conditioned reinforcers, generalized conditioned punishers are independent of specific motivating conditions and will function as punishment under most conditions.

To restate a critical point, punishers are defined by their function (a decrease in the future occurrence of behavior), not by their physical properties. Even stimulus changes that would function as unconditioned reinforcers or punishers can have the opposite effect under certain conditions. For example, a bite of food will function as a punisher for a person who has eaten too much, and an electric shock may function as a reinforcer if it signals the availability of food for a food-deprived organism (e.g., Holz & Azrin, 1961). If a student receives a smiley face sticker and praise while completing math problems and his productivity decreases as a result, smiley face stickers and praise are punishers for that student. A consequence that might serve as a punisher at work might not do so at home. Although common experiences mean that many of the same stimulus events function as conditioned punishers for most people, a punisher for one person may be a powerful reinforcer for another. Remember the cartoon described in Chapter 11 of two students cleaning erasers after school: The activity was punishment for one and reinforcement for the other.

Discriminative Effects of Punishment

Punishment does not operate in a contextual vacuum. The environmental situation in which punishment occurs plays an

important role in determining the antecedent stimulus conditions in which punishment's suppressive effects will be observed. The three-term contingency for "the operant functional relation involving punishment can be stated much like that involving reinforcement: (1) In a particular stimulus situation (S), (2) some kinds of behavior (R), when followed immediately by (3) certain stimulus changes (S^P), show a decreased future frequency of occurrence in the same or in similar situations" (Michael, 2004, p. 36).

If punishment occurs only in some stimulus conditions and not in others (e.g., a child gets scolded for reaching into the cookie jar before dinner only when an adult is in the room), the suppressive effects of punishment will be most prevalent under those conditions. A discriminated operant for punishment is the product of a conditioning history in which responses in the presence of a certain stimulus have been punished and similar responses in the absence of that stimulus have not been punished (or have resulted in a reduced frequency or intensity of punishment). Highway speeding is a discriminated operant in the repertoire of many motorists who drive within the speed limit in and around locations where they have been pulled over by the police for speeding but continue to drive over the speed limit on roads where they have never seen a police cruiser.

There is no standard term or symbol in the behavior analysis literature for an antecedent stimulus that acquires stimulus control related to punishment. Some authors have modified the shorthand symbol for a discriminative stimulus for reinforcement (S^D) to indicate the antecedent stimulus in a three-term punishment contingency, such as the S^{D-} used by Sulzer-Azaroff and Mayer (1977). Other authors refer to an antecedent stimulus correlated with the presence of a punishment contingency as a *punishment-based* S^D (e.g., Malott & Shane, 2016; Michael, 2004). We have adopted S^{Dp} as the symbol for the **discriminative stimulus for punishment**, as proposed by O'Donnell (2001). "An S^{Dp} can be defined as a stimulus condition in the presence of which a response has a lower probability of occurrence than it does in its absence as a result of response-contingent punishment delivery in the presence of the stimulus" (p. 262). Figure 14.2 shows three-term contingency diagrams for discriminated operants for positive punishment and negative punishment.

A study by McKenzie, Smith, Simmons, and Soderlund (2008) illustrates the stimulus control of a client's problem behavior by a stimulus associated with punishment. Diane, a woman with profound intellectual disabilities, had engaged in chronic eye poking that impaired her vision and on two occasions dislodged one of her eyes from the socket. Several reinforcement-based interventions had failed to produce clinically meaningful reductions in the behavior.

During an initial *no-wristbands baseline* phase, Diane was observed through a one-way mirror and the duration of her eye poking measured. In the *wristbands (no reprimands)* condition that followed, the therapist entered the room at the first occurrence of eye poking, placed two red tennis wristbands on Diane's wrists, and then left the area and remained out of Diane's sight for the remainder of the session. When eye poking was first observed during the *wristbands (reprimands)* condition, the therapist immediately entered the room and delivered a verbal reprimand (e.g., "Diane, stop poking!"). After presenting the reprimand and observing that Diane had stopped eye poking, the therapist left the room. The plan "was to repeat the reprimand at 3-s intervals until eye poking ceased, but it was never necessary to deliver more than one reprimand" (McKenzie et al., 2008, p. 257).

Data from a phase in which wristbands (reprimands) and no-wristbands (baseline) sessions were alternated showed that the wristbands rapidly acquired and maintained stimulus control

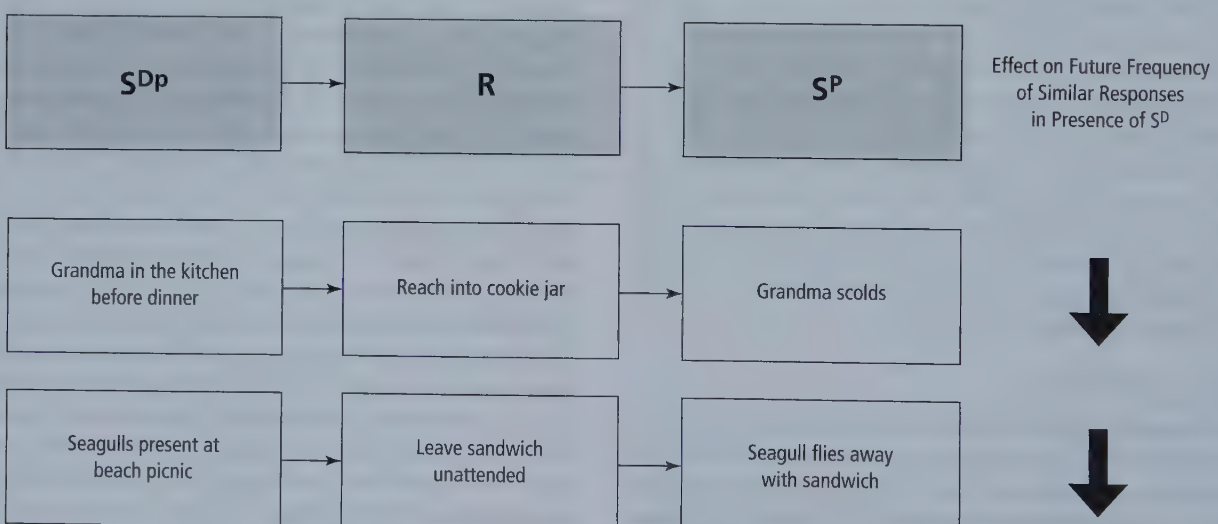


Figure 14.2 Three-term contingencies illustrating positive and negative punishment of discriminated operants: A response (R) emitted in the presence of a discriminative stimulus (S^{Dp}) is followed closely in time by a stimulus change (S^P) and results in a decreased frequency of similar responses in the future when the S^{Dp} is present. A discriminated operant for punishment is the product of a conditioning history in which responses in the presence of the S^{Dp} have been punished and similar responses in the absence of the S^{Dp} have not been punished (or have resulted in a reduced frequency or intensity of punishment than in the presence of the S^{Dp}).

over eye poking (see Figure 14.3). During wristbands (reprimands) sessions, Diane poked her eyes for a mean of 2 seconds (range, 0 to 7 sec) compared to a mean of 133 seconds (range, 2 sec to 402 sec) during no-wristbands sessions. Discriminative control of the wristbands was replicated in a delayed multiple baseline across settings analysis. Results showed reprimands in the treatment setting had acquired stimulus control. Notably, no reprimands were ever delivered when Diane wore the wristbands in the canteen or at home.

Recovery from Punishment

Punishment's suppressive effects on behavior are usually not permanent. This phenomenon, called **recovery from punishment**, is especially likely when the behavior being punished also obtains reinforcement. Figure 14.4 shows the results from one of a series of basic laboratory experiments by Azrin (1960) that demonstrated recovery from punishment. The top set of data paths is the cumulative record of a pigeon's key pecking during a 1-hr session of food reinforcement on a variable interval (VI) 6-min schedule prior to any punishment. The data paths on the lower part of the figure show responding during punishment and reinforcement. Beginning on the 1st day, and for each daily 1-hr session thereafter for 35 consecutive sessions, every key peck was followed immediately by a brief electric shock of moderate intensity (5 milliamperes for 100 milliseconds). Contingent shock suppressed pecking on the 1st day of punishment, but responding gradually recovered over the next 20 sessions and stabilized at about 900 responses per session, compared with 2800 responses prior to punishment.

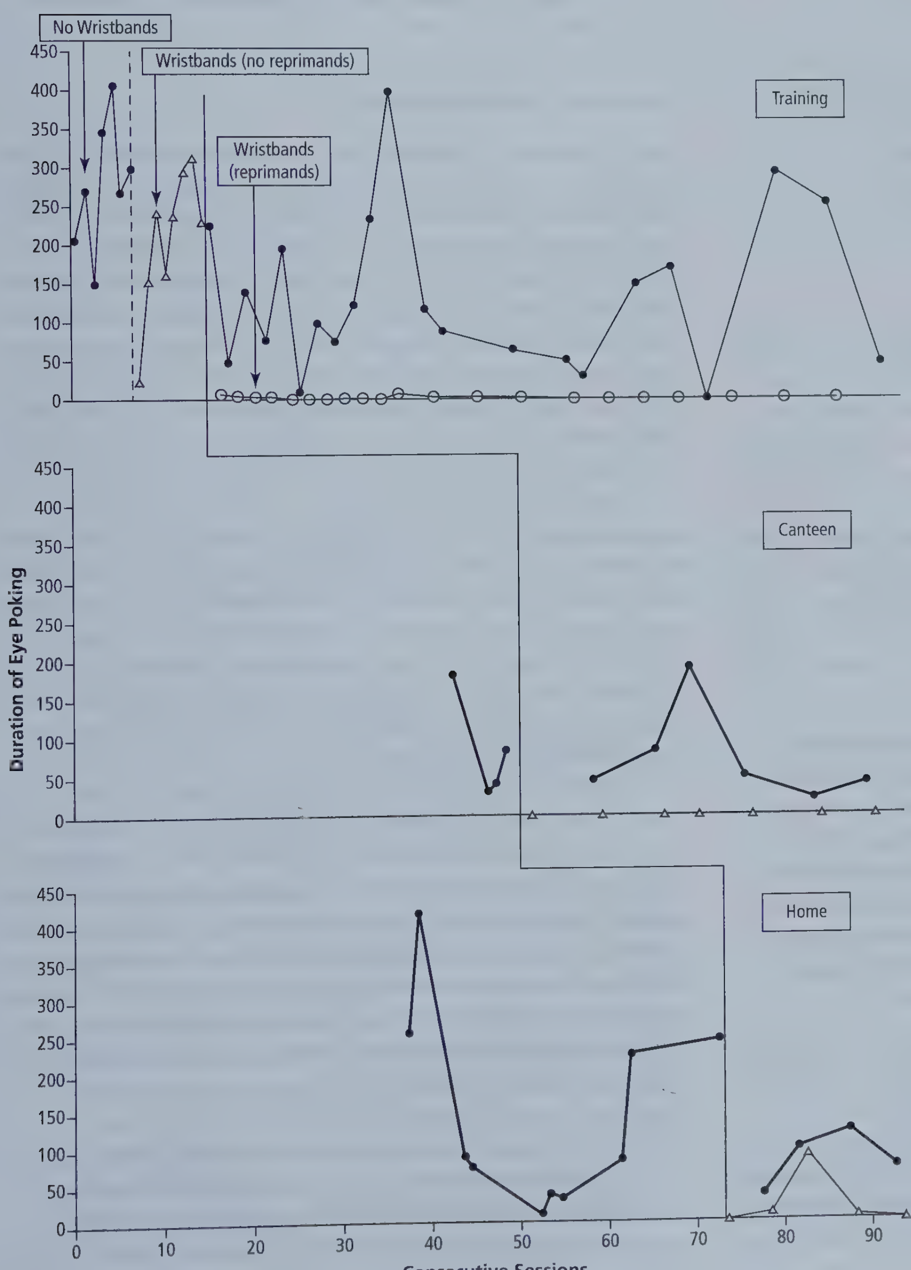


Figure 14.3 Experimental demonstration of the discriminative effects of punishment.

"Using a stimulus correlated with reprimands to suppress automatically maintained eye poking," by S. D. McKenzie, R. G., Smith, J. N. Simmons, and M. J. Soderlund, 2008, *Journal of Applied Behavior Analysis*, *41*, p. 258. Reproduced with permission of John Wiley & Sons Inc.

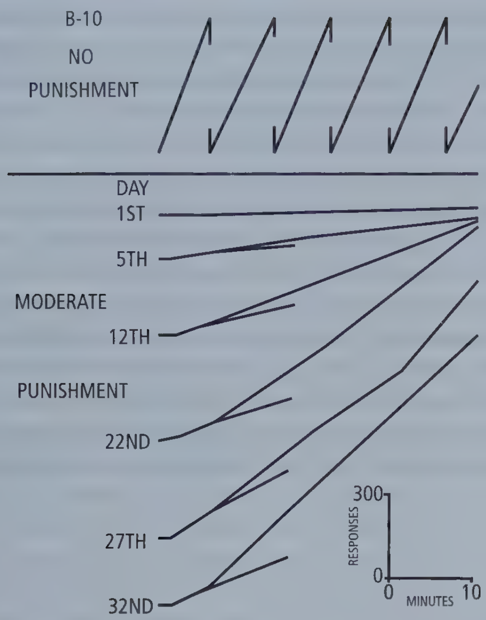


Figure 14.4 Partial recovery of responding by a pigeon during punishment of every response concurrent with reinforcement on a variable interval schedule.

"Effects of Punishment Intensity during Variable-Interval Reinforcement," By N. H. Azrin, 1960, *Journal of Experimental Analysis of Behavior, 3*, p. 127. Reproduced with permission of John Wiley & Sons Inc.

Recovery of responding to prepunished levels is more likely to occur when the punishment was mild or when the person can discriminate that the punishment contingency is no longer active. Sometimes responding after punishment is discontinued will not only recover but also briefly exceed the level at which it was occurring prior to punishment (Azrin, 1960; Holz & Azrin, 1962). Although the response-weakening effects of punishment often wane when punishment is discontinued, so, too, do the response-strengthening effects of reinforcement, which are often fleeting when previously reinforced behavior is placed on extinction (Vollmer, 2002). Michael (2004) noted that

> Recovery from punishment is sometimes given as an argument against the use of punishment to cause a decrease in behavior. "Don't use punishment because the effect is only temporary." But of course the same argument could be made for reinforcement. The strengthening effect of reinforcement decreases when the behavior occurs without the reinforcement. The weakening effect of punishment decreases when the behavior occurs without the punishment. (p. 38)

FACTORS THAT INFLUENCE THE EFFECTIVENESS OF PUNISHMENT

Reviews of basic and applied research identify several variables as keys to the effectiveness of punishment: immediacy, intensity, schedule or frequency, inadvertent reinforcement of the target behavior, and the availability of reinforcement for an alternative behavior (Axelrod, 1990; Azrin & Holz, 1966; Hineline & Rosales-Ruiz, 2013; Lerman & Vorndran, 2002).

Immediacy

Maximum suppressive effects are obtained when the onset of the punisher occurs immediately after the occurrence of the target response. Even small time delays can decrease the effectiveness of punishment.[3]

> Basic studies with humans and nonhumans indicate that punishment procedures can fail to suppress responding when the consequence is delayed by just 10–30 seconds. . . . When a punisher is delayed, other responses or multiple occurrences of the targeted behavior are likely to intervene before the consequence is delivered, weakening the contingency between the response and its consequence. (Lerman & Toole, 2011, p. 359)

Intensity of Punishment

Basic researchers examining the effects of punishers of varying intensity have reported three reliable findings: (1) A positive correlation exists between the intensity of the punishing stimulus and response suppression; (2) a negative correlation exists between the intensity of the punishing stimulus and recovery from punishment (described above and illustrated in Figure 14.4); and (3) a high-intensity stimulus that was previously presented at low intensity and gradually increased in intensity for subsequent responses may be ineffective as punishment (e.g., Azrin & Holz, 1960; Hake, Azrin, & Oxford, 1967; Holz & Azrin, 1962; Terris & Barnes, 1969).

Data from another of Azrin's (1960) early laboratory experiments on the effects of different intensities of punishment illustrate the first two findings (see Figure 14.5). Each curve is the cumulative record of a pigeon pecking an illuminated key during a 1-hr period of concurrent food reinforcement on a VI 1-min schedule and punishment of each response by shocks of various intensities. Results for two subjects show the more intense the punishing stimulus, the more immediately and thoroughly it suppresses behavior.

Data from other studies demonstrated that recovery from punishment varies inversely with the intensity of punishment (e.g., Azrin, 1960; Hake, Azrin, & Oxford, 1967). Responding by subjects that received very mild punishment recovered more rapidly and completely than did responding by subjects that received more intense punishment. Virtually permanent response suppression may occur when complete suppression of behavior to a zero rate of responding has been achieved with intense punishment. In their review of basic research on punishment, Azrin and Holz (1966) noted that:

> Intense punishment did not merely reduce responses to the unconditioned or operant level, but reduced them to an absolute level of zero. Since punishment is delivered only after a response occurs, there is

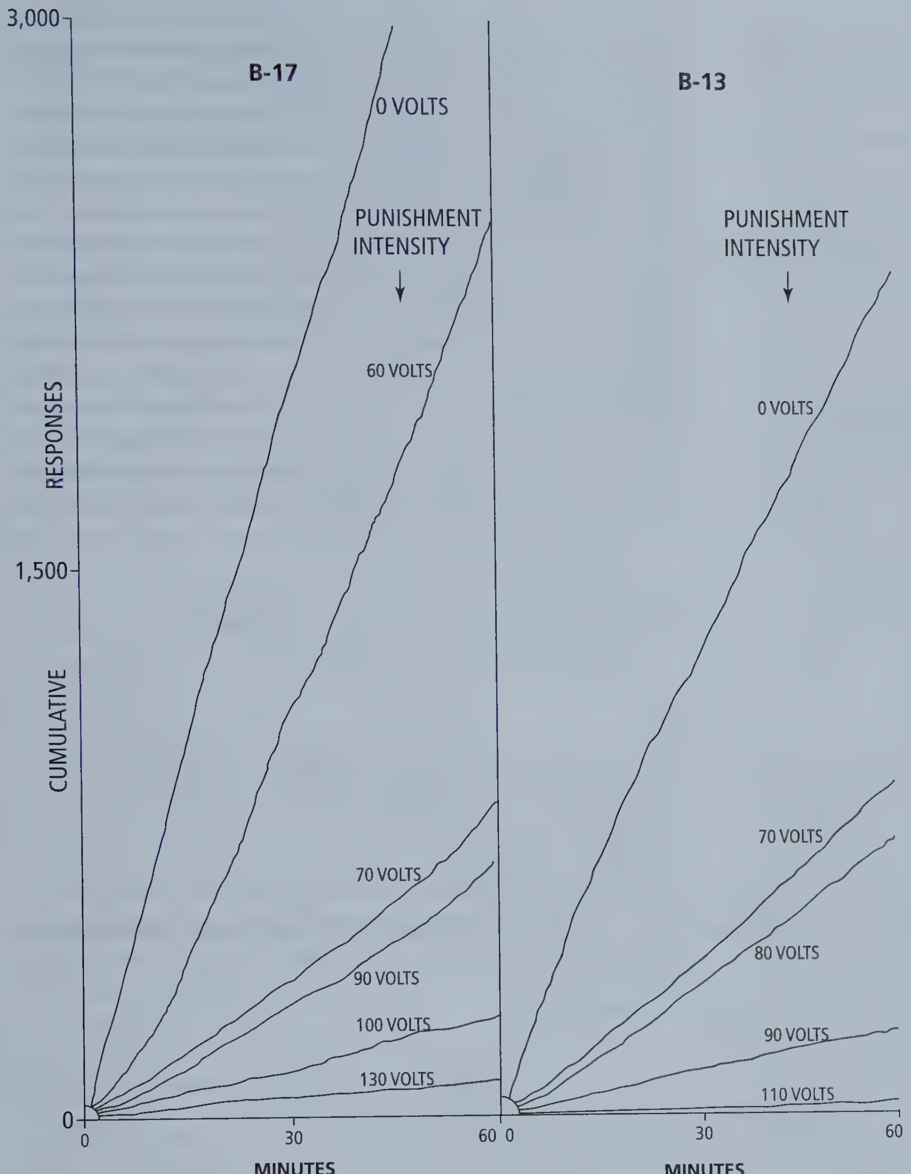


Figure 14.5 Response rate for two pigeons as a function of punishment intensity. Each curve is the cumulative record of pecking an illuminated key during a 1-hr period of concurrent reinforcement on a VI 1-min schedule and in which punishment of each response occurs.

"Effects of Punishment Intensity during Variable-Interval Reinforcement," by N. H. Azrin, 1960, *Journal of Experimental Analysis of Behavior*, *3*, p. 138. Reproduced with permission of John Wiley & Sons Inc.

> no opportunity for the subject to detect the absence of punishment unless he responds. If punishment is so severe as to completely eliminate the responses then the opportunity for detecting the absence of punishment no longer exists. (p. 410)

A high-intensity stimulus may be ineffective as a punisher if that stimulus was initially presented at low intensity and the intensity is increased gradually (Hineline & Rosales-Ruiz, 2013). Recovery of responding to pre-punished levels is more likely to rebound when (a) the punishment was mild, (b) habituation to the punisher occurred, (c) the person can discriminate that the punishment contingency is no longer active, (d) the punished behavior is reinforced inadvertently, or (e) staff training (or the lack of it) militates against treatment integrity (Lerman & Toole, 2011).

However, as Lerman and Vorndran (2002) pointed out, relatively few applied studies have examined the relation between punishment intensity and treatment efficacy, and that research has yielded inconsistent and sometimes contradictory results (Cole, Montgomery, Wilson, & Milan, 2000; Singh, Dawson, & Manning, 1981; Williams, Kirkpatrick-Sanchez, & Iwata, 1993). When selecting the intensity of a punishing stimulus, the practitioner should ask: Will this degree of the punisher suppress occurrences of the problem behavior? Lerman and Vorndran (2002) recommended, "While the punishing stimulus needs to be intensive enough for an effective application, it should not be more intense than necessary." (p. 443)

Schedule and Consistency of Punishment

The suppressive effects of a punisher are maximized by a continuous schedule of punishment (FR 1) in which each occurrence of the behavior is followed by the punishing consequence. In general, the greater the proportion of responses that are followed by the punisher, the greater the response reduction

(Azrin, Holz, & Hake, 1963; Zimmerman & Ferster, 1962). Azrin and Holz (1966) summarized the comparative effects of punishment on continuous and intermittent schedules as follows:

> Continuous punishment produces more suppression than does intermittent punishment for as long as the punishment contingency is maintained. However, after the punishment contingency has been discontinued, continuous punishment allows more rapid recovery of the responses, possibly because the absence of punishment can be more rapidly discriminated. (p. 415)

Intermittent punishment, however, may be partially effective under some conditions (Cipani, Brendlinger, McDowell, & Usher, 1991; Clark, Rowbury, Baer, & Baer, 1973; Romanczyk, 1977). Lerman, Iwata, Shore, and DeLeon (1997) demonstrated that a gradual thinning of the punishment schedule might maintain the suppressive effects of punishment that was initially delivered on a continuous schedule (FR 1). Participants were five adults with profound intellectual disabilities and histories of chronic self-injurious behavior (SIB) in the form of hand mouthing or head hitting. Treatment by punishment (time-out from reinforcement for one participant and contingent restraint for the other four) delivered on a continuous (FR 1) schedule produced marked reductions in SIB from baseline levels for all five participants. (Figure 14.6 shows the results for three of the five participants.) The participants were then exposed to intermittent schedules of punishment (FI 120-sec or FI 300-sec). On the FI 120-sec schedule, the therapist delivered punishment contingent on the first SIB response after 120 seconds had elapsed since the previous application of punishment or the start of the session. The frequency of SIB under the intermittent schedule of punishment for all but

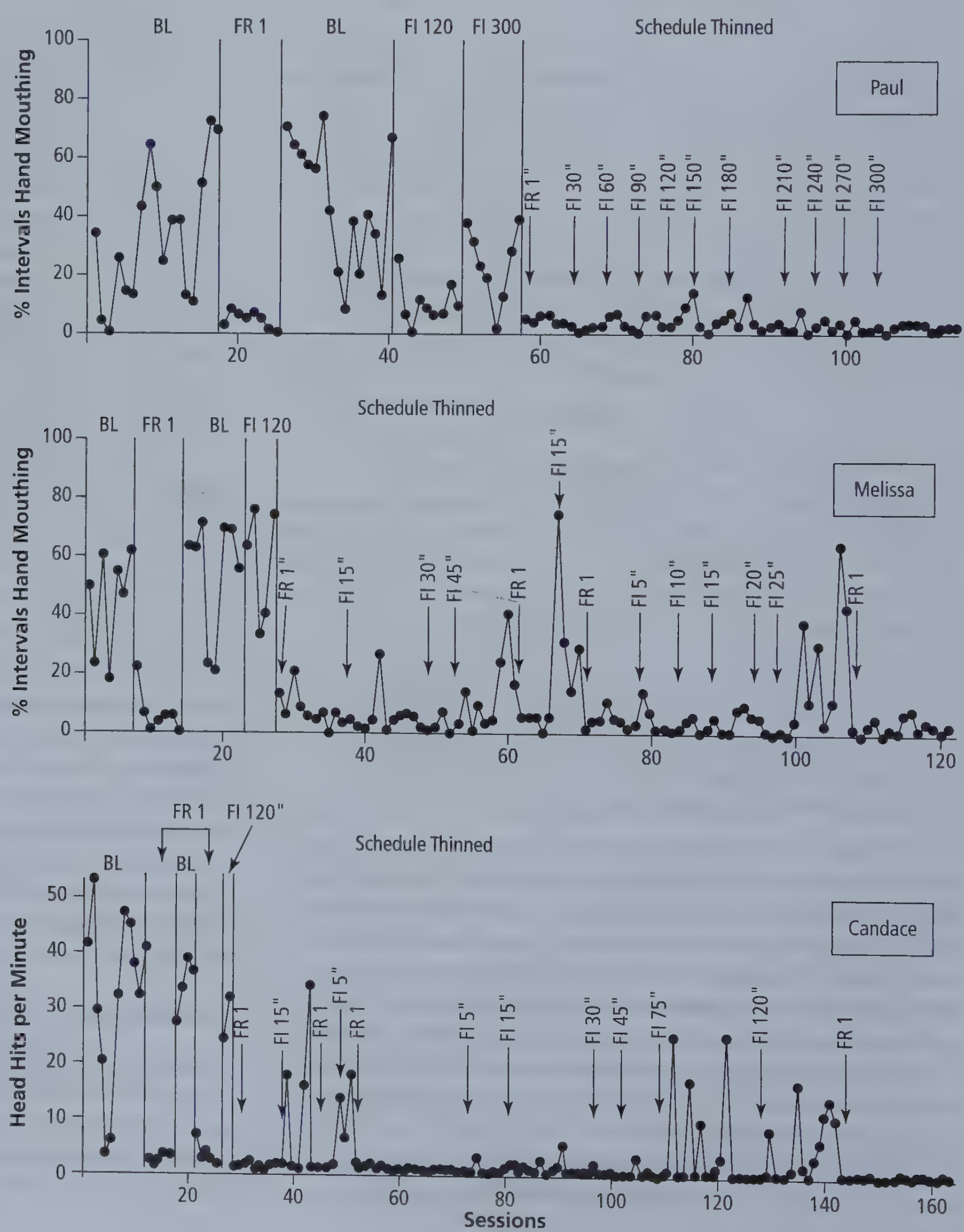


Figure 14.6 Self-injurious behavior (SIB) by three adults with profound intellectual disabilities during baseline and punishment delivered on continuous (FR 1) and various fixed interval schedules.

"Effects of Intermittent Punishment on Self-Injurious Behavior: An Evaluation of Schedule Thinning," by D. C. Lerman, B. A. Iwata, B. A. Shore, and I. G. DeLeon, 1997, *Journal of Applied Behavior Analysis, 30*, p. 194. Reproduced with permission of John Wiley & Sons Inc.

one participant (Wayne, not shown in Figure 14.6) increased to baseline levels.

After re-establishing low levels of SIB for each participant with FR 1 punishment, the researchers gradually thinned the punishment schedules. For example, schedule thinning for Paul consisted of increasing the duration of the fixed interval duration by 30-sec increments to FI 300-sec (i.e., FR 30-sec, FR 60-sec, FR 90-sec, etc.). With the exception of a few sessions, Paul's SIB remained at low levels as the punishment schedule was progressively thinned over 57 sessions. During the final 11 sessions, in which punishment was delivered on an FI 300-sec schedule, his SIB occurred at a mean of 2.4% of observed intervals (compared to 33% in baseline). A similar pattern of success with a gradually thinning punishment schedule was obtained for another subject, Wendy (not shown in Figure 14.6).

The effectiveness of punishment on an FI 300-sec schedule for three of the five participants—Paul, Wendy, and Wayne (whose SIB remained low even though punishment was abruptly changed from FR 1 to FI 120-sec and then FI 300-sec)—enabled infrequent application of punishment. In practice, this would free therapists or staff from continuous monitoring of behavior.

For Melissa and Candace, however, repeated attempts to gradually thin the punishment schedule proved unsuccessful in maintaining the frequency of their SIB at levels attained by FR 1 punishment. Lerman and colleagues (1997) speculated that one explanation for the ineffectiveness of punishment on a fixed interval schedule was that after a person has experienced an FI schedule for some time, the delivery of punishment could function "as a discriminative stimulus for punishment-free periods, leading to a gradual overall increase in responding under FI punishment" (p. 198).

Reinforcement for the Target Behavior

The effectiveness of punishment is modulated by the reinforcement contingencies maintaining the problem behavior. If a problem behavior occurs at a frequency sufficient to cause concern, it is presumably producing reinforcement. If the target response is never reinforced, then "punishment would scarcely be possible since the response would rarely occur" (Azrin & Holz, 1966, p. 433).

To the extent that the reinforcement maintaining the problem behavior can be reduced or eliminated, punishment will be more effective. Of course, if *all* reinforcement for the problem behavior was withheld—a condition difficult to achieve in many settings—the resulting extinction schedule would decrease the behavior independent of the presence of a punishment contingency. However, as Azrin and Holz (1966) pointed out:

> The physical world often provides reinforcement contingences that cannot be eliminated easily. The faster we move through space, the quicker we get to where we are going, whether the movement be walking or driving an auto. Hence, running and speeding will inevitably be reinforced. Extinction of running and speeding could be accomplished only by the impossible procedure of eliminating all reinforcing events that result from movement through space. Some other reductive method, such as punishment, must be used. (p. 433)

Reinforcement for Alternative Behaviors

Holz, Azrin, and Ayllon (1963) found that punishment was ineffective in reducing psychotic behavior when that behavior was the only means by which patients could attain reinforcement. However, when patients could emit an alternative response that resulted in reinforcement, punishment was effective in reducing their inappropriate behavior. Summing up laboratory and applied studies that have reported the same finding, Millenson (1967) stated:

> If punishment is employed in an attempt to eliminate certain behavior, then whatever reinforcement the undesirable behavior had led to must be made available via a more desirable behavior. Merely punishing school children for "misbehavior" in class may have little permanent effect. . . . The reinforcers for "misbehavior" must be analyzed and the attainment of these reinforcers perhaps permitted by means of different responses, or in other situations. . . . But for this to happen, it appears important to provide a rewarded alternative to the punished response. (p. 429)

A study by Thompson, Iwata, Conners, and Roscoe (1999) provides an excellent illustration of how the suppressive effects of punishment can be enhanced by reinforcement for an alternative response. Four adults with developmental disabilities who had been referred to a day treatment program for self-injurious behavior (SIB) participated in the study. Twenty-eight-year-old Shelly, for example, expelled saliva and then rubbed it onto her hands and other surfaces (e.g., tables, windows), which led to frequent infections; and Ricky, a 34-year-old man with deaf-blindness, frequently hit his head and body, resulting in bruises (contusions). Previous interventions such as differential reinforcement of appropriate behavior, response blocking, and protective equipment had been ineffective in reducing the SIB of all four participants.

A functional behavior analysis (see Chapter 27) with each participant suggested that SIB was maintained by automatic reinforcement. Reinforcer assessments were conducted to identify materials that produced the highest levels of contact with the lowest levels of SIB (e.g., wooden stringing beads, a mirrored microswitch that produced vibration and music, a balloon). The researchers then conducted a punisher assessment to determine the least intrusive consequences that produced a 75% reduction in SIB for each participant.

Using an experimental design that combined reversal, alternating treatments, and multiple baseline components, Thompson and colleagues (1999) analyzed the effects of punishment with and without reinforcement of alternative behavior on participants' SIB. During the no-punishment condition, the therapist was present in the room, but did not interact with the participant or provide any consequences for SIB. During the punishment condition, immediately following each occurrence of SIB the therapist delivered the consequence previously identified as a punisher for the participant. For example:

> Each time Shelly expelled saliva, the therapist delivered a reprimand ("no spitting") and briefly dried each of her

hands (and any other wet surfaces) with a cloth. Ricky's hands were held in his lap for 15 s each time he engaged in SIB. Donna and Lynn both received a verbal reprimand and had their hands held across their chests for 15 s following SIB. (p. 321)

Within each no-punishment and punishment phase, sessions of reinforcement and no-reinforcement conditions were alternated. During reinforcement sessions, participants had continual access to leisure materials or activities previously identified as highly preferred; during no-reinforcement sessions, participants had no access to leisure materials.

The study's results are shown in Figure 14.7. During the no-punishment baseline phase, only Shelly's SIB was consistently lower during the reinforcement sessions than during the no-reinforcement sessions. Although the introduction of punishment reduced SIB from baseline levels for all four participants, punishment was more effective during those sessions in which reinforcement for alternative behaviors was available. Also, fewer punishers were delivered during the punishment condition sessions when reinforcement was available.

Thompson and colleagues (1999) concluded that the effects of punishment can be improved when reinforcement is applied for alternative behaviors, and that practitioners need not increase the aversiveness of the punishing stimulus to increase the effectiveness of punishment. Adding a reinforcement component to the punishment procedure may be sufficient to achieve desired results, while simultaneously implementing a less restrictive approach.

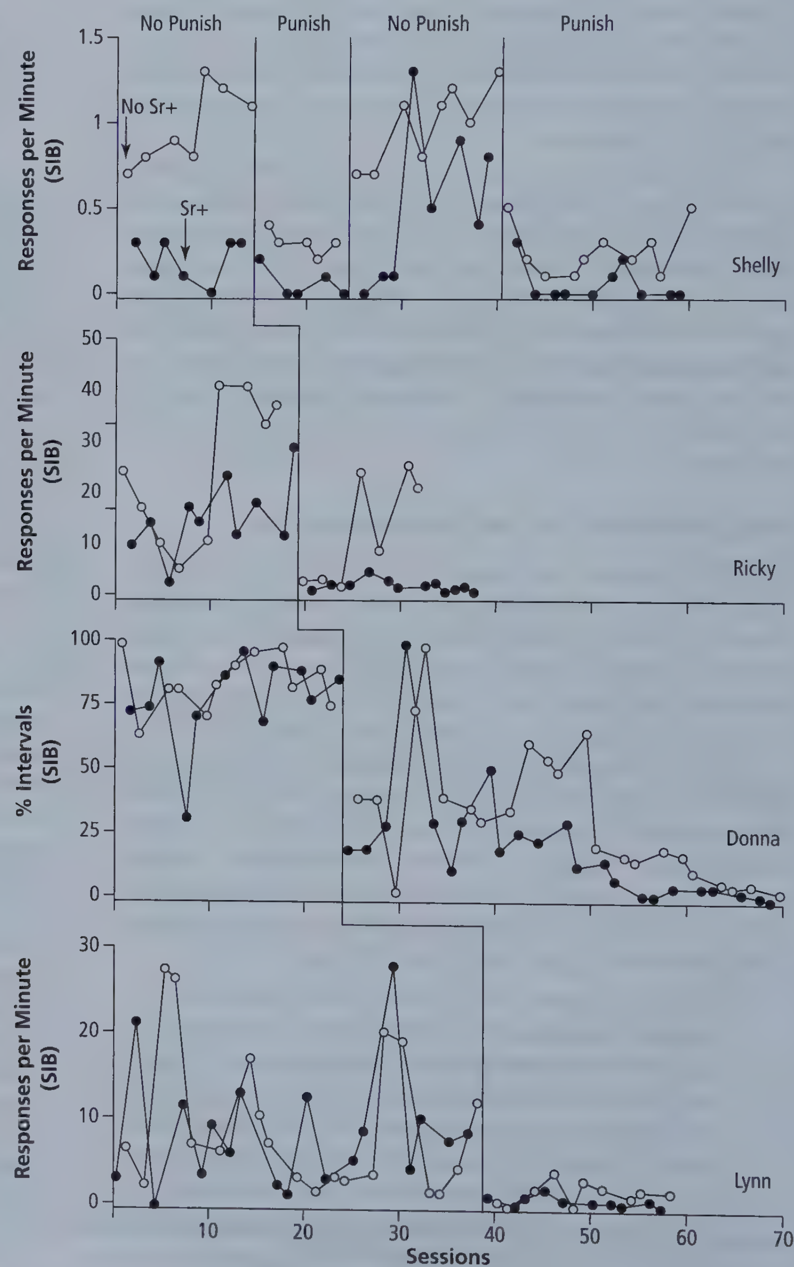


Figure 14.7 Self-injurious behavior by four adults with developmental disabilities during alternating reinforcement and no-reinforcement conditions across no-punishment and punishment phases.

"Effects of Reinforcement for Alternative Behavior during Punishment for Self-Injury," by R. H. Thompson, B. A. Iwata, J. Conners, and E. M. Roscoe, 1999, *Journal of Applied Behavior Analysis*, *32*, p. 323. Reproduced with permission of John Wiley & Sons Inc.

POSSIBLE SIDE EFFECTS AND PROBLEMS WITH PUNISHMENT

A variety of side effects are often correlated with punishment, including the elicitation of undesirable emotional responses and aggression, escape and avoidance, and an increased rate of the problem behavior under nonpunishment conditions (Azrin & Holz, 1966; Hutchinson, 1977; Linscheid & Meinhold, 1990). Other problems noted include modeling undesirable behavior and overusing punishment because reductions in problem behavior can function as negative reinforcement for caregivers.

Emotional and Aggressive Reactions

Punishment, especially positive punishment in the form of aversive stimulation, may evoke aggressive behavior with respondent and operant components (Azrin & Holz, 1966). For example, electric shock elicited reflexive forms of aggression and fighting in laboratory animals (Azrin, Hutchinson, & Hake, 1963; Ulrich & Azrin, 1962; Ulrich, Wolff, & Azrin, 1962). Such pain-elicited, or *respondent aggression*, is directed toward any nearby person or object. A student who is punished severely may begin to throw and destroy materials within her reach, and attack other students or the person delivering the punishment. Aggressive behavior following punishment that occurs because it has enabled the person to escape the aversive stimulation in the past is referred to as *operant aggression* (Azrin & Holz, 1966).

Escape and Avoidance

Escape and avoidance are expected reactions to aversive stimulation. Escape and avoidance behaviors take a wide variety of forms, some of which may be a greater problem than the target behavior being punished. For example, a student who is admonished repeatedly for sloppy work or coming to class unprepared may stop coming to class altogether. A person may lie, cheat, hide, or exhibit other undesirable behaviors to avoid punishment. Mayer, Sulzer, and Cody (1968) indicated that escape and avoidance do not always occur in the literal sense of those terms. Individuals may escape punishing environments by consuming drugs or alcohol, or by simply "tuning out."

As the intensity of a punisher increases, so, too, does the likelihood of escape and avoidance. For example, in a study evaluating the effectiveness of a specially designed cigarette holder that delivered an electric shock to the user when it was opened as an intervention for reducing cigarette smoking, Powell and Azrin (1968) found the following: "As the punishment intensity increased, the duration decreased for which the subjects would remain in contact with the contingency; ultimately, an intensity was reached at which they refused to experience it altogether" (p. 69).

Escape and avoidance as side effects to punishment, like emotional and aggressive reactions, can be minimized or prevented altogether by providing desirable alternative responses to the problem behavior that both avoid the delivery of punishment and provide reinforcement.

Behavioral Contrast

Reynolds (1961) introduced the term **behavioral contrast** to refer to the phenomenon in which a change in one component of a multiple schedule that increases or decreases the rate of responding on that component is accompanied by a change in the response rate in the opposite direction on the other, unaltered component of the schedule.[4] Behavioral contrast can occur as a function of a change in reinforcement or punishment density on one component of a multiple schedule (Brethower & Reynolds, 1962; Lattal & Griffin, 1972). For example, behavioral contrast for punishment takes the following general form: (a) Responses are occurring at similar rates on two components of a multiple schedule (e.g., a pigeon pecks a backlit key, which alternates between blue and green, reinforcement is delivered on the same schedule on both keys, and the bird pecks at roughly the same rate regardless of the key's color); (b) responses on one component of the schedule are punished, whereas responses on the other component continue to go unpunished (e.g., pecks on the blue key are punished, and pecks on the green key continue to produce reinforcement at the prior rate); (c) rate of responding decreases on the punished component and increases on the unpunished component (e.g., pecks on the blue key are suppressed and pecks on the green key increase even though pecks on the green key produce no more reinforcement than before).

Here is a hypothetical example to further illustrate the contrast effect of punishment. A child is eating cookies before dinner from the kitchen cookie jar at equal rates in the presence and absence of his grandmother. One day, Grandma scolds the child for eating a cookie before dinner, which suppresses his rate of pre-dinner cookie eating when she is in the kitchen (see Figure 14.2); but when Grandma's not in the kitchen, the boy eats cookies from the jar at a higher rate than he did when unsupervised prior to punishment. Contrast effects of punishment can be minimized, or prevented altogether, by consistently punishing occurrences of the target behavior in all relevant settings and stimulus conditions, withholding or at least minimizing the person's access to reinforcement for the target behavior, and providing alternative desirable behaviors. (With respect to our hypothetical case of the child eating cookies before dinner, we recommend simply removing the cookie jar!)

Punishment May Involve Modeling of Undesirable Behavior

Most readers are familiar with the example of the parent who, while spanking a child, says, "This will teach you not to hit your playmates!" Unfortunately, the child may be more likely to imitate the parent's actions, not the parent's words. More than three decades of research have found a strong correlation between young children's exposure to harsh, excessive punishment and antisocial behavior and conduct disorders in adolescents and adults (Patterson, 1982; Patterson, Reid, & Dishion, 1992; Sprague & Walker, 2000). Although the appropriate use of behavior change tactics based on the punishment principle need not involve harsh treatment or negative personal interactions, practitioners should heed Bandura's (1969) valuable counsel in this regard:

> Anyone attempting to control specific troublesome responses should avoid modeling punitive forms of behavior that not only counteract the effects of direct training but also increase the probability that on future occasions the individual may respond to interpersonal thwarting in an imitative manner. (p. 313)

Punishment May Be Overused due to Negative Reinforcement of the Punishing Agent's Behavior

Negative reinforcement may be a reason for the widespread use of, and reliance upon, punishment in child rearing, education, and society. When Person A delivers a reprimand or other aversive consequence to Person B for misbehaving, the immediate effect is often the temporary cessation of the troubling behavior. The termination of this aversive stimulus serves as negative reinforcement for Person A's behavior. Reese (1966) stated the case succinctly: "Punishment reinforces the punisher" (p. 37).

Alber and Heward (2000) described how the natural contingencies at work in a typical classroom could strengthen a teacher's use of reprimands for disruptive behavior while undermining her use of contingent praise and attention for appropriate behavior.

> Paying attention to students when they are behaving inappropriately (e.g., "Carlos, you need to sit down right now!") is negatively reinforced by the immediate cessation of the inappropriate behavior (e.g., Carlos stops running around and returns to his seat). As a result, the teacher is more likely to attend to student disruptions in the future. . . . Although few teachers must be taught to reprimand students for misbehavior, many teachers need help increasing the frequency with which they praise student accomplishments. Teacher-praising behavior is usually not reinforced as effectively as teacher-reprimanding behavior. Praising a student for appropriate behavior usually produces no immediate effects—the child continues to do his work when praised. Although praising a student for working productively on an assignment may increase the future likelihood of that behavior, there are no immediate consequences for the teacher. By contrast, reprimanding a student often produces an immediate improvement in the teacher's world (if only temporary)—that functions as effective negative reinforcement for reprimanding. (pp. 178–179)

POSITIVE PUNISHMENT INTERVENTIONS

Interventions featuring positive punishment interventions can take a wide variety of forms. We describe five in this section: reprimands, response blocking, response interruption and redirection (RIRD), contingent exercise, and overcorrection.

Reprimands

It may seem strange that immediately following a discussion of teachers' overreliance on reprimands, our first example of a positive punishment intervention focuses on reprimands. The delivery of verbal reprimands following the occurrence of misbehavior is perhaps the most common form of *attempted* positive punishment. However, a number of studies have shown that a firmly said "No!" or "Stop! Don't do that!" delivered immediately on the occurrence of a behavior can function as effective punishment (e.g., Hall et al., 1971; Thompson et al., 1999; Van Houten, Nau, Mackenzie-Keating, Sameoto, & Colavecchia, 1982).

In spite of the widespread use of reprimands in an effort to suppress undesired behavior, surprisingly few studies have examined their effectiveness as punishers. Results of a series of experiments by Van Houten, Nau, Mackenzie-Keating, Sameoto, and Colavecchia (1982) designed to identify variables that increased the effectiveness of reprimands as punishers for disruptive behavior in the classroom revealed that reprimands delivered with eye contact and in close proximity to the student were more effective than reprimands delivered from across the room. An interesting study by O'Leary, Kaufman, Kass, and Drabman (1970) found that quiet reprimands that were audible only to the child being reprimanded were more effective in reducing disruptive behavior than loud reprimands that could be heard by other children in the classroom.

If being reprimanded is the only way a child receives adult attention, it should not be surprising that reprimands function as reinforcement for that child rather than as punishment. Indeed, Madsen, Becker, Thomas, Koser, and Plager (1968) found that the repeated use of reprimands while students were out of seat served to increase, rather than reduce, the behavior. Consistent with research on other punishing stimuli, reprimands are more effective as punishers when motivation for the problem behavior has been minimized and the availability of an alternative behavior maximized (Van Houten & Doleys, 1983).

A parent or teacher does not want to adopt a pattern of constant reprimanding. Reprimands should be used thoughtfully and sparingly in combination with frequent praise and attention contingent on appropriate behavior. O'Leary and colleagues (1970) recommended:

> An ideal combination would probably be frequent praise, some soft reprimands, and very occasional loud reprimands. . . . Combined with praise, soft reprimands might be very helpful in reducing disruptive behaviors. In contrast, it appears that loud reprimands lead one into a vicious cycle of more and more reprimands resulting in even more disruptive behavior. (p. 155)

Response Blocking

Response blocking—physically intervening as soon as the person begins to emit the problem behavior, to prevent the completion of the response—has shown effectiveness in reducing the frequency of some problem behaviors, such as chronic hand mouthing, eye poking, and pica (e.g., Lalli, Livezey, & Kates, 1996; Lerman & Iwata, 1996; Reid, Parsons, Phillips, & Green, 1993). In addition to preventing the response from occurring by using the least amount of physical contact and restraint possible, the therapist might issue a verbal reprimand or prompt to stop engaging in the behavior (e.g., Hagopian & Adelinis, 2001).

Lerman and Iwata (1996) used response blocking in treating the chronic hand mouthing (contact between any part of the

hand and the lips or mouth) of Paul, a 32-year-old man with profound intellectual disabilities. Following a baseline condition, in which Paul was seated in a chair with no one interacting with him and no leisure materials were available, response blocking on an FR 1 schedule was implemented. A therapist sat behind Paul and blocked his attempts to put his hand in his mouth. "Paul was not prevented from bringing his hand to his mouth; however, the therapist blocked the hand from entering the mouth by placing the palm of her hand about 2 cm in front of Paul's mouth" (p. 232). Response blocking produced an immediate and rapid decrease in hand-mouthing attempts to near-zero levels (see Figure 14.8).

Response blocking is often implemented as a treatment for SIB or self-stimulatory behavior when functional analysis reveals consistent responding in the absence of socially mediated consequences, which suggests the possibility that the behavior is maintained by automatic reinforcement from sensory stimuli produced by the response. Because response blocking prevents the learner from contacting the sensory stimuli that are normally produced by completing the response, subsequent decreases in responding could be due to extinction. Lerman and Iwata (1996) presented their study as a potential method for distinguishing whether the suppressive effects of response blocking are due to punishment or extinction mechanisms. They explained their reasoning, in part, as follows:

> Depending on the mechanism through which behavior is reduced (extinction vs. punishment), different schedules of reinforcement or punishment are in effect when a given proportion of responses is blocked. . . . Thus, as larger proportions of responses are blocked, the reinforcement schedule becomes leaner and the punishment schedule becomes richer. If response blocking produces extinction, response rates should increase or be maintained as more responses are blocked (i.e., as the reinforcement schedule is thinned), until [the effects of] extinction [i.e., reduced response rate] occurs at some point along the progression. Conversely, if the procedure functions as punishment, response rates should decrease as more responses are blocked (i.e., as the punishment schedule becomes richer). (pp. 231–232, words in brackets added)

A condition in which all responses are blocked might function as an extinction schedule (i.e., reinforcement in the form of sensory stimuli is withheld for all responses) or as a continuous (FR 1) schedule of punishment (i.e., all responses are followed by physical contact). As Lerman and Iwata (1996) explained, if only some responses are blocked, the situation may function as an intermittent schedule of reinforcement or as an intermittent schedule of punishment. Therefore, comparing response rates among conditions in which different proportions of responses are blocked should indicate whether the effects are due to extinction or to punishment.

If response blocking functioned as extinction for Paul's hand mouthing, an initial increase in response rate would be expected when the blocking procedure was implemented for every response; however, no such increase was observed.[5] If response blocking functioned as punishment, blocking every response would constitute a continuous schedule of punishment and a rapid decrease in responding would be expected; and that is exactly what the results showed (see data for the first response block [1.0] phase in Figure 14.8).

Conversely, if response blocking functioned as extinction for Paul's hand mouthing, then blocking some, but not all, responses would place the hand mouthing on an intermittent schedule of reinforcement and responding would be expected to increase from baseline levels. And blocking an ever-larger proportion of responses thins the reinforcement schedule further, causing the response rate to rise even higher. Instead, as a greater proportion of responses were blocked, the suppressive effects on Paul's SIB became more pronounced, a result expected as a punishment schedule becomes denser. Overall, therefore, the results of the experiment indicated that response blocking functioned as punishment for Paul's hand mouthing.

Smith, Russo, and Le (1999) conducted a systematic replication of Lerman and Iwata's (1996) experiment and found that the frequency of eye poking by a 41-year-old woman treated

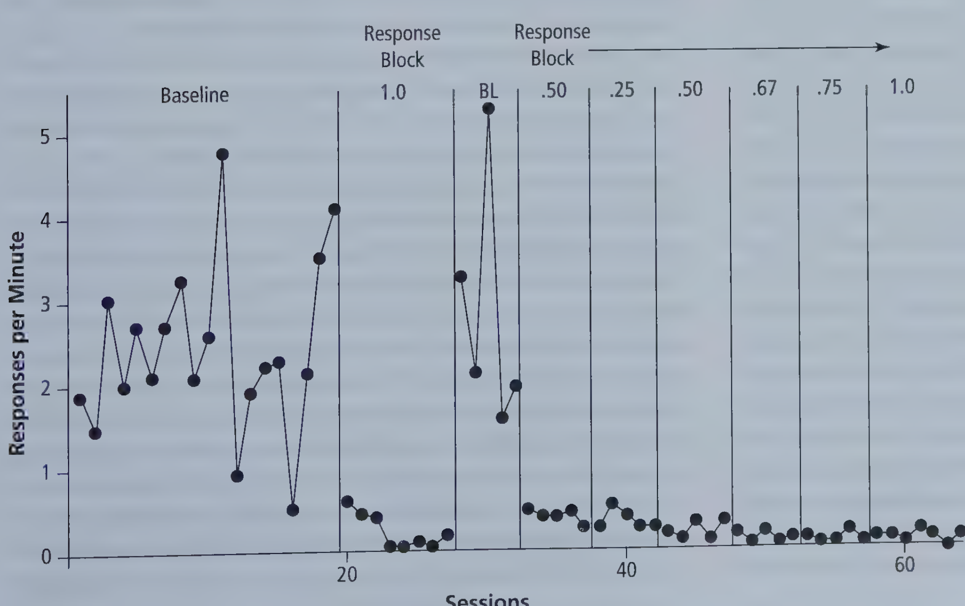


Figure 14.8 Rates of hand mouthing during baseline and varying schedules of response blocking.

"A Methodology for Distinguishing between Extinction and Punishment Effects Associated with Response Blocking," by D. C. Lerman and B. A. Iwata, 1996, *Journal of Applied Behavior Analysis, 29*, p. 232. Reproduced with permission of John Wiley & Sons Inc.

with response blocking decreased gradually—a response pattern indicative of extinction. The authors concluded that "whereas blocking may reduce one participant's behavior via punishment, it may extinguish another participant's behavior" (p. 369).

If response blocking occurs *before* the response is emitted, its suppressive effects cannot be due to punishment or extinction. Punishment and extinction are both behavior → consequence relations. Any procedure that prevents the target behavior from occurring in its entirety eliminates the behavior → consequence contingency and should be considered an antecedent intervention. However, if the problem behavior is a progression of movements or a response chain, then blocking the initial movement or response in the chain before the sequence produces the presumed reinforcer would alter the behavior → consequence contingency. For example, blocking a client's open hand *after* he begins to move it toward contact with his head is a consequence whose suppressive effects can be analyzed in terms of punishment and/or extinction.

Although response blocking may be viewed as a less restrictive and more humane intervention than delivering aversive stimulation after a response has occurred, it must be approached with great care. Side effects such as aggression and resistance to the response-blocking procedure have occurred (Hagopian & Adelinis, 2001; Lerman, Kelley, Vorndran, & Van Camp, 2003). Providing prompts and reinforcement for an alternative response can minimize resistance and aggression. For example, the aggressive behavior by a 26-year-old man with moderate intellectual disabilities and bipolar disorder during treatment with response blocking for pica (ingesting paper, pencils, paint chips, and human feces) was reduced by supplementing response blocking with a prompt and redirection to engage in an alternative behavior, in this case moving to an area of the room where popcorn was available (Hagopian & Adelinis, 2001).

Response Interruption and Redirection (RIRD)

Response interruption and redirection (RIRD) is a procedural variation of response blocking that includes interrupting stereotypic behavior at its onset and redirecting the individual to complete high-probability behaviors instead. Two types of RIRD have been identified and evaluated in the literature: motor RIRD and vocal RIRD (Ahearn, Clark, MacDonald, and Chung, 2007; Cassella, Sidener, Sidener, & Progar, 2011; Shawler & Miguel, 2015; Wunderlich & Vollmer, 2015). Both types have been used primarily with individuals with autism who emitted high levels of repetitive, stereotypic, nonfunctional behaviors. Procedurally, RIRD is employed when, at the onset of a stereotypic behavior, the therapist immediately presents a series of instructions, or demands, requiring high-probability vocal or motor responses (e.g., "What am I holding?" "Tell me your name." "Clap your hands."). RIRD is usually terminated upon a set criterion of successful responses (e.g., three correct in a row), although Saini, Gregory, Uran, and Fantetti (2015) found that one correct demand response was effective. RIRD has been used alone and combined with other procedures to reduce stereotypy so that students could engage in more interpersonal or socially appropriate behaviors. Practitioners who consider using RIRD procedures should have sufficient time, resources, and trained staff before implementation. RIRD can be labor- and time-intensive, and may extend teaching sessions beyond what can be accommodated in some educational or therapeutic settings.

Contingent Exercise

Contingent exercise is an intervention in which the client is required to perform a response that is not topographically related to the problem behavior. Contingent exercise has been found effective as punishment for various self-stimulatory, stereotypic, disruptive, aggressive, and self-injurious behaviors (e.g., DeCatanzaro & Baldwin, 1978; Kern, Koegel, & Dunlap, 1984; Luce & Hall, 1981; Luiselli, 1984).[6] In perhaps the most frequently cited example of contingent exercise as a punisher, Luce, Delquadri, and Hall (1980) found that the repetition of mild exercise contingent on aggressive behavior by two boys with severe disabilities reduced it to near-zero levels. Figure 14.9 shows the results for Ben, a 7-year-old who frequently hit other children at school. Each time Ben hit someone, he was required to stand up and sit down 10 times. Initially, Ben had to be physically prompted to stand; an assistant held the child's hand while pulling his upper body forward. Physical prompts were accompanied by verbal prompts of "stand up" and "sit down." Soon, whenever hitting occurred, the nearest adult simply said, "Ben, no hitting. Stand up and sit down 10 times" and the verbal prompts alone were sufficient for Ben to complete the exercise. If a hitting episode occurred during the contingent exercise, the procedure was reinstated.

Overcorrection

Overcorrection is a behavior reduction tactic in which, contingent on each occurrence of the problem behavior, the learner is required to engage in effortful behavior that is directly or logically related to the problem behavior. Originally developed by Foxx and Azrin (1972, 1973; Foxx & Bechtel, 1983) as a method for decreasing disruptive and maladaptive behaviors of adults with intellectual disabilities in institutional settings, overcorrection combines the suppressive effects of punishment and the educative effects of positive practice. Overcorrection includes either or both of two components: restitution and positive practice.

In **restitutional overcorrection**, contingent on the problem behavior, the learner is required to repair the damage caused by the problem behavior by returning the environment to its original state and then to engage in additional behavior that brings the environment to a condition better than it was prior to the misbehavior. A parent applying restitutional overcorrection with a child who repeatedly tracks mud onto the kitchen floor might require the child to first wipe up the mud and clean his shoes and then to *over* correct the effects of his misbehavior by mopping and waxing a portion of the floor and polishing his shoes.

Azrin and Foxx (1971) used restitutional overcorrection in their toilet training program by requiring a person who had an accident to undress, wash her clothes, hang them up to dry, shower, dress in clean clothing, and then clean up a portion of the lavatory. Azrin and Wesolowki (1974) eliminated stealing of food by hospitalized adults with intellectual disabilities by requiring residents to return not only the stolen food, or the

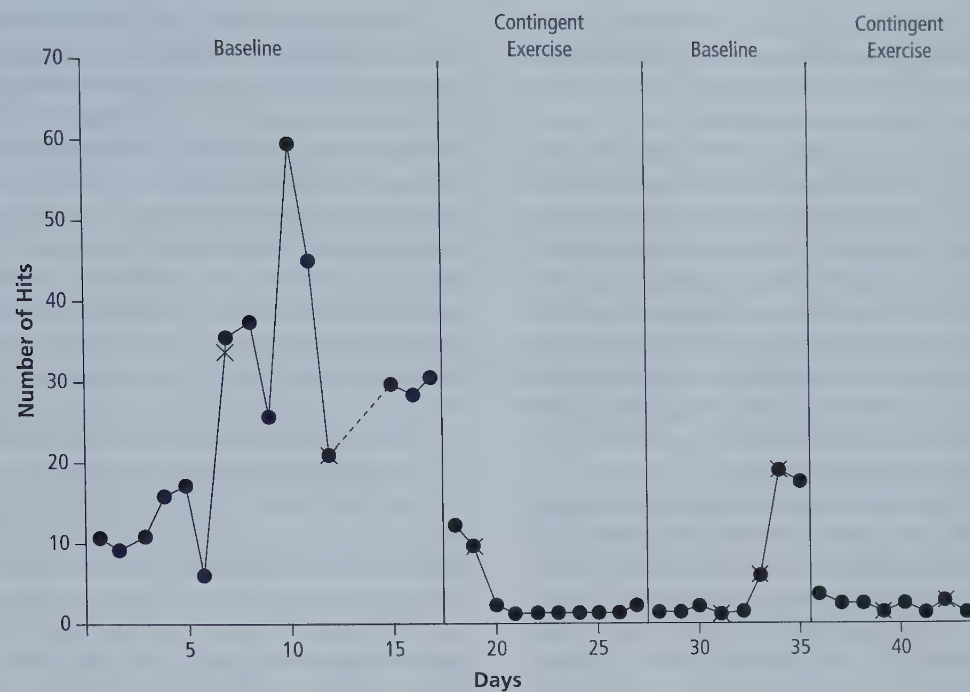


Figure 14.9 Number of times a 7-year-old boy hit other children during 6-hour school days during baseline and contingent exercise. X's represent response measures recorded by a second observer.

"Contingent Exercise: A Mild but Powerful Procedure for Suppressing Inappropriate Verbal and Aggressive Behavior," by S. C. Luce, J. Delquadri, and R. V. Hall, 1980, *Journal of Applied Behavior Analysis, 13*, p. 587. Reproduced with permission of John Wiley & Sons Inc.

portion that remained uneaten, but also to purchase an additional item of that food at the commissary and give it to the victim.

Azrin and Besalel (1999) differentiated a procedure they called *simple correction* from overcorrection. With simple correction, the learner is required, subsequent to the occurrence of an inappropriate behavior, to restore the environment to its previous state. For example, a simple correction procedure is in effect when requiring a student who cuts to the head of the lunch line to go to the back of the line. Requiring the student to wait until all others get in line and are served before re-entering the line would constitute a form of overcorrection in this instance. Azrin and Besalel recommended that simple correction be used to reduce behaviors that are not severe, occur infrequently, are not deliberate, and do not severely interfere with or annoy other people.

Correction is not possible if the problem behavior produces an irreversible effect (e.g., a one-of-a-kind dish is broken) or if the corrective behavior is beyond the person's means or skills. In such instances, Azrin and Besalel (1999) recommended that the person be required to correct as much of the damage his behavior caused as possible, be present at all points in the correction, and assist with any parts of the correction that he is able to perform. For example, a child who breaks a neighbor's window might be required to clean up the pieces of glass, measure the window, contact the store for a replacement pane, and be present and assist with each step when the new window pane is installed.

In **positive practice overcorrection**, contingent on an occurrence of the problem behavior, the learner is required to repeatedly perform a correct form of the behavior, or a behavior incompatible with the problem behavior, for a specified duration of time or number of responses. Positive practice overcorrection entails an educative component in that it requires the person to engage in an appropriate alternative behavior. The parent whose son tracks mud into the house could add a positive practice component by requiring him to practice wiping his feet on the outside door mat and then to enter the house for 2 minutes or 5 consecutive times. Overcorrection that includes restitution and positive practice helps teach what to do in addition to what not to do. The child who breaks an irreplaceable dish could be required to gently and slowly wash a number of dishes, perhaps with an exaggerated carefulness.

Researchers and practitioners have used positive practice overcorrection to reduce the frequency of problem behaviors such as toilet training (Azrin & Foxx, 1971), self-stimulation and stereotypic behavior (Azrin, Kaplan, & Foxx, 1973; Foxx & Azrin, 1973b), pica (Singh & Winton, 1985), bruxism (Steuart, 1993), sibling aggression (Adams & Kelley, 1992), and classroom disruptions (Azrin & Powers, 1975). Positive practice overcorrection has also been used for academic behaviors (Lenz, Singh, & Hewett, 1991), most often to decrease oral reading and spelling errors (e.g., Ollendick, Matson, Esveldt-Dawson, & Shapiro, 1980; Singh & Singh, 1986; Singh, Singh, & Winton, 1984; Stewart & Singh, 1986).

Positive practice overcorrection can also be applied to reduce or eliminate behaviors that do not create permanent response products that can be repaired or restored to their original state. For example, Heward, Dardig, and Rossett (1979) described how parents used positive practice overcorrection to

help their teenage daughter stop making a grammatical error in her speech. Eunice frequently used the contraction "don't" instead of "doesn't" with the third person singular (e.g., "He don't want to go."). A positive reinforcement program in which Eunice earned points she could redeem for preferred activities each time she used "doesn't" correctly had little effect on her speech. Eunice agreed with her parents that she should speak correctly but claimed the behavior was a habit. Eunice and her parents then decided to supplement the reinforcement program with a mild punishment procedure. Each time Eunice or her parents caught her using "don't" incorrectly in her speech, she was required to say the complete sentence she had just spoken 10 times in a row, using correct grammar. Eunice wore a wrist counter that reminded her to listen to her speech and to keep track of the number of times she employed the positive practice procedure.

When positive practice effectively suppresses the problem behavior, it is not clear what behavioral mechanisms are responsible for the behavior change. Punishment may result in decreased frequency of responding, because the person engages in effortful behavior as a consequence of the problem behavior. A reduction in the frequency of the problem behavior as an outcome of positive practice may also be a function of an increased frequency of an incompatible behavior, the correct behavior that is strengthened in the person's repertoire as the result of the intensive, repeated practice. Azrin and Besalel (1999) suggested that positive practice may be effective because it provides additional positive learning opportunities that have been absent in prior environments, and also requires "extra effort" that discourages future misbehaviors.

Specific procedures for implementing overcorrection vary greatly. Factors such as the type and severity of the problem behavior, its effects on the environment, the setting, the desired alternative behavior, and the learner's current skills should be considered. Figure 14.10 provides general guidelines for implementing overcorrection (Azrin & Besalel, 1999; Foxx & Bechtel, 1983; Kazdin, 2001; Miltenberger & Fuqua, 1981).

Although rapid and long-lasting therapeutic effects have been reported after corrective training, practitioners should be aware of several potential problems and limitations associated with overcorrection. First, overcorrection is a labor-intensive, time-consuming procedure that requires the full attention of the practitioner implementing it. Implementing overcorrection usually requires the practitioner to monitor the learner directly throughout the overcorrection process. Second, for overcorrection to be effective as punishment, the time that the learner spends with the person monitoring the overcorrection sequence must not be reinforcing. "If it is, it just might be worth waxing the entire kitchen floor if Mom chats with you and provides a milk and cookies break" (Heward et al., 1979, p. 63).

Third, a child who misbehaves frequently may not execute a long list of "cleanup" behaviors just because he was told to do so. Azrin and Besalel (1999) recommended three strategies to minimize the likelihood of refusal to perform the overcorrection sequence: (1) remind the learner what the more severe disciplinary action will be and, if the refusal persists, then impose that discipline; (2) discuss the need for correction before the problem behavior occurs; and (3) establish correction as an expectation and routine habit for any disruptive behavior. If the child resists too strongly or becomes aggressive, overcorrection may not be a viable treatment. Adult learners must voluntarily make the decision to perform the overcorrection routine.

GUIDELINES FOR USING PUNISHMENT

Assuming that habituation, treatment integrity, reinforcement, and training issues can be addressed, punishment can yield rapid, long-lasting suppression of problem behavior. Unfortunately, however, agency policies, human subject review procedures, and historical practices have limited the use of punishment for research and treatment in clinic-based settings (Grace, Kahng, & Fisher, 1994). Still, punishment may be a treatment of choice when (a) the problem behavior produces serious physical harm

Figure 14.10 Guidelines for implementing overcorrection.

1. Immediately upon the occurrence of the problem behavior (or the discovery of its effects), in a calm, unemotional tone of voice, tell the learner that he has misbehaved and provide a brief explanation for why the behavior must be corrected. Do not criticize or scold. Overcorrection entails a logically related consequence to reduce future occurrences of the problem behavior; criticism and scolding do not enhance the tactic's effectiveness and may harm the relationship between the learner and practitioner.
2. Provide explicit verbal instructions describing the overcorrection sequence the learner must perform.
3. Implement the overcorrection sequence as soon as possible after the problem behavior has occurred. When circumstances prevent immediately commencing the overcorrection sequence, tell the learner when the overcorrection process will be conducted. Several studies have found that overcorrection conducted at a later time can be effective (Azrin & Powers, 1975; Barton & Osborne, 1978).
4. Monitor the learner throughout the overcorrection activity. Provide the minimal number of response prompts, including gentle physical guidance, needed to ensure that the learner performs the entire overcorrection sequence.
5. Provide the learner with minimal feedback for correct responses. Do not give too much praise and attention to the learner during the overcorrection sequence.
6. Provide praise, attention, and perhaps other forms of reinforcement to the learner each time he "spontaneously" performs the appropriate behavior during typical activities. (Although technically not part of the overcorrection procedure, reinforcing an alternative behavior is a recommended complement to all punishment-based interventions.)

and must be suppressed quickly, (b) reinforcement-based treatments have not reduced the problem behavior to socially acceptable levels, or (c) the reinforcer maintaining the problem behavior cannot be identified or withheld (Lerman & Vorndran, 2002).

If the decision is made to use a punishment-based intervention, steps should be taken to ensure that the intended punisher actually works. The guidelines that follow will help practitioners apply punishment with optimal effectiveness while minimizing undesirable side effects and problems. We based these guidelines on the assumption that the analyst has conducted a functional behavior assessment to identify the variables that maintain the problem behavior, that the problem behavior has been defined clearly, and that the participant cannot avoid or escape the punisher.

Select Effective and Appropriate Punishers

Conduct Punisher Assessments

Aversive stimulus assessments are conducted to identify stimuli that function as punishers. Fisher and colleagues (1994) identified two advantages of conducting punisher assessments. First, the sooner an ideal punisher can be identified, the sooner it can be applied to the problem behavior. Second, data from punisher assessments might reveal the magnitude or intensity of the punisher necessary for behavioral suppression, enabling the practitioner to deliver the lowest intensity punisher to achieve socially significant response suppression.

Punisher assessment mirrors reinforcement assessment, except that instead of measuring engagement or duration of contact with each stimulus, the behavior analyst measures overall response suppression, negative verbalizations, avoidance movements, escape attempts, or disruptions associated with each potential punishing stimulus. Data from the punisher assessments are then used to develop a hypothesis on the relative comparative effects of each stimulus change as a punisher.

The decision of which among several potential punishers to choose should be based on the relative degree of punisher intrusiveness, the ease with which the punisher can be consistently and safely delivered, and data on potential response suppression. Subsequent observation might reveal that a consequence that is less intrusive, time-consuming, or difficult to apply might produce a significant decrease in the targeted behavior (Thompson et al., 1999).

Use Varied Punishers of Sufficient Quality and Intensity

The quality of a punisher is relative to a number of past and current variables that affect the participant. For example, Thompson and colleagues found that 15 seconds of physical restraint was a high-quality punisher for Ricky, Donna, and Lynn, but the physical restraint was not as ideal for Shelly. Although stimuli that reliably evoke escape and avoidance behaviors often function as high-quality punishers, practitioners should recognize that a stimulus change that effectively suppresses some behaviors may not affect other behaviors and that highly motivated problem behaviors may only be suppressed by a particularly high-quality punisher.

Generally, basic and applied researchers have found that the greater the intensity, magnitude, or amount of a punishing stimulus, the greater the suppression of behavior. This finding is conditional on delivering the punishing stimulus at its optimum level of magnitude initially, rather than gradually increasing the level over time (Azrin & Holtz, 1966). For example, Thompson and colleagues (1999) used the previously described punisher assessment procedure to determine an optimum magnitude or duration level for a punishing stimulus: a level and duration that produced a 75%, or greater, decrease in self-injurious behaviors (SIB) from baseline. Specifically, they determined that to meet the criterion of a punishing stimulus, one participant had his hands held in his lap for 15 seconds when SIB occurred, whereas two other participants received a verbal reprimand and their hands were held against their body for 15 seconds following SIB.

Beginning with a punishing stimulus of sufficient intensity is important, because participants may adapt to the punishing stimulus when levels of magnitude are increased gradually. It is possible that 15 seconds of movement restriction would have been a poor punisher for the participants if Thompson et al. (1999) had begun with a punishing stimulus of 3 seconds and gradually increased its magnitude from 3- to 6- to 9- to 15-sec intervals. Hineline and Rosales-Ruiz (2013) advise practitioners to begin "with an intense but safe, practical, and accepted punisher and then withhold further increases in intensity until the other variables are ruled out" (p. 488).

Use Varied Punishers

The effectiveness of a punishing stimulus can decrease with repeated presentations of that stimulus. Using a variety of punishers may help to reduce habituation effects. In addition, using various punishers may increase the effectiveness of less intrusive punishers. For example, Charlop, Burgio, Iwata, and Ivancic (1988) compared various punishers to a single presentation of one of the punishers (i.e., a stern "No!", overcorrection, time-out with physical restraint, a loud noise). Three-, 5-, and 6-year-old children with developmental disabilities served as participants. Their problem behaviors included aggression (Child 1), self-stimulation and destructive behavior (Child 2), and aggression and out-of-seat behavior (Child 3). The varied-punisher condition was slightly more effective than the single-presentation condition and enhanced the sensitivity of the behavior to less intrusive punishing stimuli (see Figure 14.11). Charlop and colleagues concluded, "It appears that by presenting a varied format of commonly used punishers, inappropriate behaviors may further decrease without the use of more intrusive punishment procedures" (p. 94).

Deliver the Punisher at the Beginning of a Behavioral Sequence

Punishing an inappropriate behavior as soon as it begins is more effective than waiting until the chain of behavior has been completed (Solomon, 1964). Once the sequence of responses that make up the problem behavior is initiated, powerful secondary reinforcers associated with completing each step of the chain may prompt its continuation, thereby counteracting the inhibiting or suppressing effects of the punishment that occurs at the end of the sequence. Therefore, whenever practical, the punishing stimulus should be presented early in the behavioral

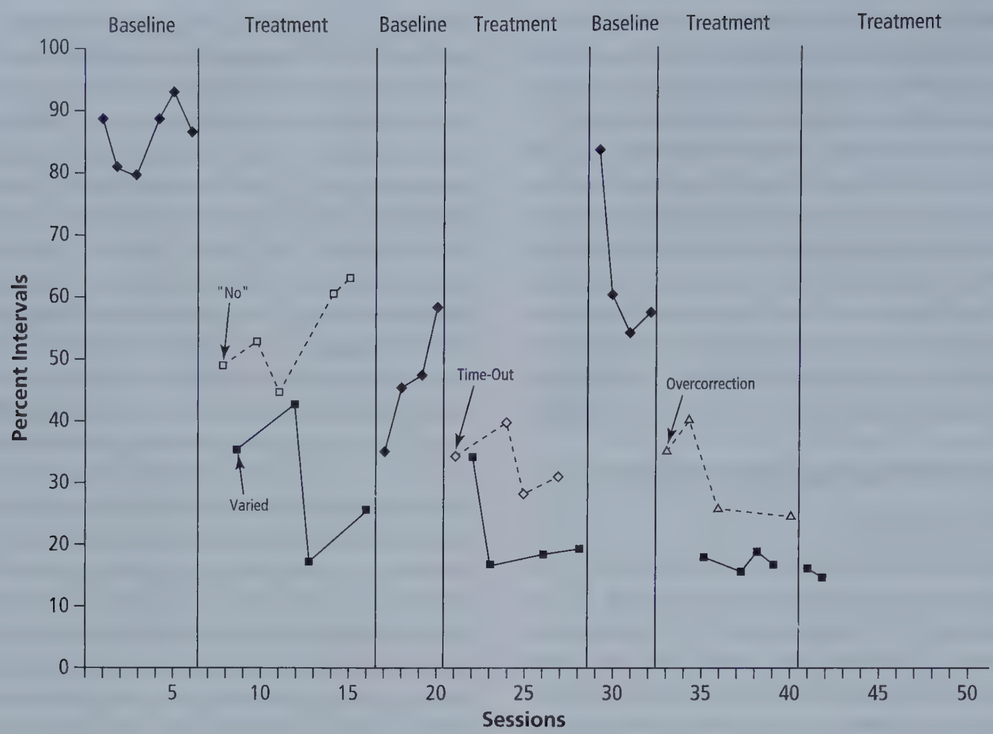


Figure 14.11 Percentage of intervals of occurrence of self-stimulation and destructive behavior for a 6-year-old girl with autism.

"Stimulus Variation as a Means of Enhancing Punishment Effects," by M. H. Charlop, L. D. Burgio, B. A. Iwata, and M. T. Ivancic, 1988, *Journal of Applied Behavior Analysis, 21*, p. 92. Reproduced with permission of John Wiley & Sons Inc.

sequence rather than later (Heron, 1978). For example, if violent arm swinging is a reliable precursor to self-injurious eye poking, then punishment (e.g., response blocking, restraint) should be delivered as soon as arm swinging starts.

Punish Each Instance of the Behavior Initially

Punishment is most effective when the punisher follows each response. This consistency is especially important when a punishment intervention is first implemented. Consistency goes beyond just programming an effective punishment schedule, that is, *when* punishment is delivered. Consistency refers to whether the punisher itself, its application as a behavior → consequence functional relation, and its procedural elements are being implemented as intended. Consistency, therefore, relates to treatment integrity. One can easily imagine the situation in which the disruptive behavior of an adolescent high school student with multiple teachers is supposedly being addressed with a plan whereby contingent on a disruptive behavior, each teacher will point his or her finger toward the student and emit a sharp reprimand, "No!". However, in reality some of the teachers follow the plan, but others do not. Or, a few teachers say "Noooo" with a semi-joking intonation, while others issue a firm reprimand. When these inconsistent situations occur, treatment integrity is at risk. Consequently, it becomes difficult to determine if the failure to suppress the disruptive behavior is the result of a stimulus that does not serve as a punisher, or if the punishment procedure is compromised by one or more treatment implementation failures.

Gradually Shift to an Intermittent Schedule of Punishment

Although punishment is most effective when the punisher immediately follows each occurrence of the problem behavior (Hineline & Rosales-Ruiz, 2013), practitioners may find a continuous schedule of punishment impractical because they lack the resources and time to attend to each occurrence of the behavior (O'Brien & Karsh, 1990). Several studies have found that, after responding has been reduced by a continuous schedule of punishment, an intermittent schedule of punishment may be sufficient to maintain the behavior at a socially acceptable frequency (Clark, Rowbury, Baer, & Baer, 1973; Lerman et al., 1997; Romanczyk, 1977).

We recommend two guidelines for using intermittent punishment. First—and this is especially important—a continuous (FR 1) schedule of punishment should be used to diminish the problem behavior to a clinically acceptable level before gradually thinning to an intermittent schedule of punishment. Second, combine intermittent punishment with extinction. It is unlikely that reduced responding will continue under intermittent punishment if the reinforcer that maintains the problem behavior cannot be identified and withheld. If these two guidelines for intermittent punishment are met and the frequency of the problem behavior increases to an unacceptable level, return to a continuous schedule of punishment and then, after recovery of acceptably low rates of responding, gradually shift to an intermittent schedule of punishment that is denser than the one used previously (e.g., VR 2 rather than VR 4).

Mediate Unavoidable Response-to-Punishment Delay

Consequences intended to serve as punishers for problem behavior are frequently delayed in the natural environment. Realistically, delays may occur because caregivers are not able to monitor behaviors continuously and may miss them, the individual may escape the delivery of the punisher, or the misbehavior may go undetected (Lerman & Vorndran, 2002). When punishers cannot be delivered immediately, applied behavior analysts should strive to mediate response-to-punishment delay.

For example, before administering a delayed punisher, Rolider and Van Houten (1985) played audiotapes of the child's tantrums earlier in the day. Explaining the tape recorder's purpose to the child and leaving it in sight while the recordings were collected might have resulted in the tape recorder bridging the temporal gap between inappropriate behavior and its consequence (i.e., by functioning as discriminative stimuli for punishment) (Lerman & Vorndran, 2002).

Supplement Punishment with Complementary Interventions

Applied behavior analysts typically do not use punishment as a single intervention: They combine punishment with other interventions, primarily differential reinforcement, extinction, and a variety of antecedent interventions. Basic and applied researchers consistently find that the effectiveness of punishment is enhanced when the learner can make other responses for reinforcement. In most circumstances, applied behavior analysts should incorporate differential reinforcement of alternative behavior (DRA), differential reinforcement of incompatible behavior (DRI), or differential reinforcement of other behaviors (DRO) (see Chapter 25) into a treatment program to supplement a punishment. When used as a reductive procedure for problem behavior, differential reinforcement consists of two components: (1) providing reinforcement contingent on the occurrence of a behavior other than the problem behavior and (2) withholding reinforcement for the problem behavior (i.e., extinction). The study by Thompson and colleagues (1999) presented previously provides an excellent example of how reinforcement of an alternative behavior can enhance the suppressive effects of punishment and enable relatively benign punishment procedures to be effective even for chronic problem behaviors that have been resistant to change.

We recommend that practitioners reinforce alternative behaviors copiously. Additionally, the more reinforcement the learner obtains by emitting appropriate behaviors, the less motivated he will be to emit the problem behavior. In other words, heavy and consistent doses of reinforcement for alternative behaviors function as an abolishing operation that weakens (abates) the frequency of the problem behavior.

There is still another important reason for recommending the reinforcement of alternative behaviors. Applied behavior analysts are in the business of building repertoires by teaching the clients and students they serve new skills and more effective ways of controlling their environments and achieving success. Punishment (with the exception of some overcorrection procedures) eliminates behaviors from a person's repertoire. Although the person will be better off without those behaviors in her repertoire, punishment only teaches her what *not* to do—it does not teach her what to do instead.

The effectiveness of antecedent interventions such as functional communication training (FCT), high-probability request sequence (high-*p*), and noncontingent reinforcement (NCR), which diminish the frequency of problem behaviors by decreasing the effectiveness of the reinforcers that maintain the problem behaviors (see Chapter 26), can be made more effective when combined with punishment. For instance, Fisher and colleagues (1993) found that, although FCT did not reduce to clinically significant levels the destructive behaviors of four participants with severe intellectual disabilities and communication deficits, a combination of FCT and punishment produced the largest and most consistent reductions in problem behaviors.

Watch and Prepare for Negative Side Effects

It is difficult to predict the side effects that may result from punishment. The suppression of one problem behavior by punishment may lead to an increase in other undesirable behaviors. For example, punishment of self-injurious behavior may produce increased levels of noncompliance or aggression. Punishing one problem behavior may lead to a parallel decrease in desirable behaviors. For example, requiring that a student rewrite an ill-conceived paragraph may result in his ceasing to produce any academic work. Although punishment may produce no undesirable side effects, practitioners should be alert to problems such as escape and avoidance, emotional outbursts, and behavioral contrast, and have a plan for dealing with such events should they occur.

Record, Graph, and Evaluate Data Daily

Data collection in the initial sessions of a punishment-based intervention is especially critical. Unlike some behavior reduction procedures whose suppressive effects are often gradual (e.g., extinction, differential reinforcement of alternative behavior), the suppressive effects of punishment are usually quick and abrupt. In their classic review of the basic research on punishment, Azrin and Holz (1966) wrote:

> Virtually all studies of punishment have been in complete agreement that the reduction of responses by punishment is immediate if the punishment is at all effective. When the data have been presented in terms of the number of responses per day, the responses have been drastically reduced or eliminated on the very first day in which punishment was administered. When the data have been presented in terms of moment-to-moment changes, the reduction of responses has resulted within the first few deliveries of punishment . . . or within a few minutes. (p. 411)

The first data points from the punishment conditions in all of the graphs presented in this chapter provide additional empirical evidence for Azrin and Holz's statement of the immediate effects of punishment. Because of the abrupt effect of punishment, the practitioner should pay particular attention to the data from the first session or two of intervention. If a noticeable reduction of the problem behavior has not occurred within two

sessions of a punishment-based intervention, we recommend that the practitioner make adjustments to the intervention.

Frequent inspection of the data from a punishment perspective reminds those involved of the purpose of the intervention and reveals whether the problem behavior is being reduced or eliminated as intended. When the data indicate that a clinically or socially significant change has occurred and has been maintained, punishment can be shifted to an intermittent schedule or perhaps terminated altogether.

ETHICAL CONSIDERATIONS REGARDING THE USE OF PUNISHMENT

Ethical considerations regarding the use of punishment revolve around three major issues: the client's right to safe and humane treatment, the professional's responsibility to use least restrictive procedures, and the client's right to effective treatment.[7]

Right to Safe and Humane Treatment

Dating from the Hippocratic Oath (Hippocrates, 460 BC–377 BC), the first ethical canon and responsibility for any human services provider is to do no harm. So, any behavior change program, be it a punishment-based intervention to reduce life-threatening self-injurious behavior or an application of positive reinforcement to teach a new academic skill, must be physically safe for all involved and contain no elements that are injurious, degrading, or disrespectful to participants.

Treatments are deemed safe when they put neither the caregiver nor the target individual at physical, psychological, or social risk (Favell & McGimsey, 1993). Although there exists no universally accepted definition of what constitutes humane treatment, a reasonable case could be made that humane treatments are (a) designed for therapeutic effectiveness, (b) delivered in a compassionate and caring manner, (c) assessed formatively to determine effectiveness and terminated if effectiveness is not demonstrated, and (d) sensitive and responsive to the overall physical, psychological, and social needs of the person.

Least Restrictive Alternative

A second canon of ethics for human services professionals is not to intrude on a client's life any more than necessary to provide effective intervention. The *doctrine of the least restrictive alternative* holds that less intrusive procedures should be tried and found to be ineffective before more intrusive procedures are implemented. Interventions can be viewed as falling along a continuum of restrictiveness from least to most. The more a treatment procedure affects a person's life or independence, such as her ability to go about daily activities in her normal environment, the greater its restrictiveness. A completely unrestrictive intervention is a logical fallacy. Any treatment must affect the person's life in some way to qualify as an intervention. At the other end of the continuum, absolute restrictiveness exists during solitary confinement, where personal independence is unavailable. All behavior change interventions developed and implemented by applied behavior analysts fall within these extremes.

Selecting any punishment-based intervention essentially rules out as ineffective all positive or positive reductive approaches based on their demonstrated inability to improve the behavior. For example, Gaylord-Ross (1980) proposed a decision-making model for reducing aberrant behavior that would have practitioners rule out assessment considerations, inappropriate or ineffective schedules of reinforcement, ecological variables, and curriculum modifications *before* punishment is implemented.

Some authors and professional organizations have advanced the position that all punishment-based procedures are inherently intrusive and should never be used (e.g., LaVigna & Donnellen, 1986; Mudford, 1995). Others have advanced the counter-position that punishment-based procedures, because of their inherent level of intrusiveness, should be used only as a last resort (Gaylord-Ross, 1980; Iwata, 1988).

Most people would rate interventions based on positive reinforcement as less restrictive than interventions based on negative reinforcement; reinforcement interventions as less restrictive than punishment interventions; and interventions using negative punishment as less restrictive than those using positive punishment. However, the intrusiveness or restrictiveness of an intervention cannot be determined by the principles of behavior on which it is based. Restrictiveness is a relative concept that depends on the procedural details and ultimately rests at the level of the person with whom it is applied. A positive reinforcement intervention that requires deprivation may be more restrictive than a positive punishment procedure in which a buzzer sounds each time an incorrect response is made, and what one person considers intrusive may pose no discomfort for another person. Horner (1990) suggested that most people accept punishment interventions to reduce challenging behavior as long as those interventions: "(a) do not involve the delivery of physical pain, (b) do not produce effects that require medical attention, and (c) are subjectively judged to be within the typical norm of how people in our society should treat each other" (pp. 166–167).

In their review of response-reduction interventions based on the principle of punishment, Friman and Poling (1995) pointed out that punishment-based tactics, such as those requiring the person to make an effortful response contingent on the occurrence of target behavior, meet Horner's (1990) criteria for acceptable response-reduction interventions. They stated: "None of the procedures caused pain or required medical attention, nor did they contrast with societal norms. For example, coaches often require disobedient players to run laps, and drill instructors make misbehaving recruits do pushups" (p. 585).

Although least restrictive alternative practices assume that the less intrusive procedures are tried and found ineffective before a more restrictive intervention is introduced, practitioners must balance that approach against an effectiveness standard. Gast and Wolery (1987) suggested that, if a choice must be made between a less intrusive but ineffective procedure and a more intrusive but effective procedure, the latter should be chosen. A Task Force on the Right to Effective Treatment appointed by the Association for Behavior Analysis International provided, in part, the following perspective on the importance of judging

the ultimate restrictiveness of treatment options on the basis of their degree of proven effectiveness:

> Indeed, a slow-acting but nonrestrictive procedure could be considered highly restrictive if prolonged treatment increases risk, significantly inhibits or prevents participation in needed training programs, delays entry into a more optimal social or living environment, or leads to adaptation and the eventual use of a more restrictive procedure. Thus, in some cases, a client's right to effective treatment may dictate the immediate use of quicker-acting, but temporarily more restrictive, procedures. (Van Houten et al., 1988, p. 114).

Right to Effective Treatment

Ethical discussions regarding the use of punishment revolve most often around its possible side effects and how experiencing the punisher may cause unnecessary pain and possible psychological harm for the person. Although each of these concerns deserves careful consideration, the right to effective treatment raises an equally important ethical issue, especially for persons who experience chronic, life-threatening severe problem behaviors. Some maintain that failing to use a punishment procedure that research has shown to be effective in suppressing self-destructive behavior similar to their client's is unethical because doing so withholds a potentially effective treatment and risks maintaining a dangerous or uncomfortable state for the person. For example, Baer (1971) stated: "[Punishment] is a legitimate therapeutic technique that is justifiable and commendable when it relieves persons of the even greater punishments that result from their own habitual behavior" (p. 111).

As Iwata (1988) explained, if all other less intrusive treatments with a sufficient research base to provide a legitimate chance of success have failed, the use of punishment procedures is the only ethical option:

> In the ideal world, treatment failures do not occur. But in the actual world failures do occur, in spite of our best efforts, leaving us with these options: continued occurrence of the behavior problem toward some devastating endpoint, restraint, sedating drugs, or aversive contingencies. I predict that if we apply our skills to maximize the effectiveness of positive reinforcement programs, we will succeed often. After following such a course, my further prediction is that if we reach the point of having to decide among these ultimate default options, the client's advocate, the parent, or, if necessary, the courts will select or use the option of aversive contingencies. Why? Because under such circumstances it is the only ethical action. (pp. 152–153)

Balancing the least restrictive alternative, the right to effective treatment, and Baer's (1971) vision to relieve a person of greater punishments if his severe behavior problem remains untreated or ineffective treatments are allowed to continue, is a difficult challenge. However, behavior analysts are not without relief. In the intervening 40-plus years since the 1980s, and in light of more recent successful developments using least restrictive alternative interventions informed by functional behavior analysis, even the most serious life-threatening behaviors can be treated effectively and humanely without resorting to the use of noxious or painful unconditioned punishers (e.g., water mist, aromatic ammonia, electric shock). Research has shown that the judicious use of stimulus control procedures (McKenzie et al., 2008), time-out and contingent restraint (Lerman et al., 1997), adding a significant reinforcement component (DRA) to a punishment regimen (Thompson et al., 1999), effectively manipulating reprimands (McKenzie et al., 2008), response blocking (Lalli, Lizevey, & Kates, 1996), contingent exercise (Luce et al., 1980), varying commonly used punishers (Charlop et al., 1988), and overcorrection (Steuart, 1993) can reduce problem behaviors for which unconditioned punishers were often applied. The application of unconditioned punishers no longer meets the standards of least restrictive alternative or best practice.

Applied behavior analysts who conduct functional behavior analyses, perform comprehensive punisher assessments, vary antecedent and consequence methods, and apply these methods alone or in combination will achieve as much, or more, success in reducing self-injurious and other dangerous behaviors than previously considered possible. In doing so, they will fulfill their mission of improving clients' quality of life by ensuring the right to effective treatment, providing best practice, and using the least restrictive alternative.

Punishment Policy and Procedural Safeguards

Armed with knowledge from the experimental literature, real-world variables and contingencies with which to contend (i.e., chronic, life-threatening problems), and practices and procedures rooted in ethical codes of conduct, practitioners can consider punishment approaches, when necessary, to provide meaningful programs for the persons in their care. One mechanism for ensuring that best practice approaches are used (Peters & Heron, 1993) is to adopt and use dynamic policies and procedures that provide clear guidelines and safeguards to practitioners.

Agencies can protect and ensure their clients' rights to safe, humane, least restrictive, and effective treatment by developing policy and procedural guidelines that must be followed when any punishment-based intervention is implemented (Favell & McGimsey, 1993; Griffith, 1983; Wood & Braaten, 1983). Figure 14.12 provides an outline and examples of the kinds of components that might be included in such a document.

Best practice dictates that practitioners consult their local, state, or national professional association policy statements regarding the use of punishment. For example, the Association for the Advancement of Behavior Therapy (AABT) provides guidelines that address treatment selection, including punishment (Favell et al., 1982). The Association for Behavior Analysis International adheres to the American Psychological Association's Ethics Code (2004), which, in turn, addresses treatment issues. Collectively, policy statements address implementation requirements, procedural guidelines and precautions, and evaluation methods that an agency should use when implementing any punishment-based intervention. Chapter 31 provides additional information on ethical standards for applied

Figure 14.12 Suggested components in an agency's policy and procedural guidelines to help ensure the ethical, safe, and effective use of punishment.

Policy Requirements

- Intervention must conform to all local, state, and federal statutes.
- Intervention must conform to the policies and codes of ethical conduct of relevant professional organizations.
- Intervention should include procedures for strengthening and teaching alternative behaviors.
- Intervention must include plans for generalization and maintenance of behavior change and criteria for eventual termination or reduction of punishment.
- Informed consent must be obtained from the client or a parent or legal advocate before intervention begins.

Procedural Safeguards

- Before intervention begins, all relevant staff must be trained in (a) the technical details of properly administering the punishment procedure; (b) procedures for ensuring the physical safety of the client and staff and the humane treatment of the client; and (c) what to do in the event of negative side effects, such as emotional outbursts, escape and avoidance aggression, and noncompliance.
- Supervision and feedback of staff administering the punishment intervention must be provided and, if necessary, booster-training sessions provided.

Evaluation Requirements

- Each occurrence of the problem behavior must be observed and recorded.
- Each delivery of the punisher must be recorded and the client's reactions noted.
- Periodic review of the data (e.g., daily, weekly) by a team of parent/advocates, staff, and technical consultants must be conducted to ensure that ineffective treatment is not prolonged or effective treatment terminated.
- Social validity data must be obtained from the client, significant others, and staff on (a) treatment acceptability and (b) the real and likely impact of any behavior change on the client's current circumstances and future prospects.

behavior analysts and procedures for securing informed consent, protecting clients' rights, and ensuring safe and humane interventions.

CONCLUDING PERSPECTIVES

We conclude this chapter with brief comments on our perspectives regarding the principle of punishment, and especially applications featuring unconditioned punishers as default technologies. We believe that applied behavior analysis would be a stronger, more competent discipline, and its practitioners would be more effective, if (a) punishment's natural role and contributions to survival and learning are recognized and valued within their limitations, (b) focused basic and applied research on punishment is conducted, and (c) treatments featuring differential reinforcement options, stimulus control procedures, time-out from positive reinforcement, response blocking, and contingent restraint, be employed in lieu of treatments featuring unconditioned punishers as default technologies.

Punishment's Natural and Necessary Role in Survival and Learning Should Be Recognized

Behavior analysts ought not dismiss punishment out of hand as a possible therapeutic intervention. Positive and negative punishment contingencies occur naturally in everyday life as part of a complex mix of concurrent reinforcement and punishment contingencies, as Baum (2017) illustrated so well in this example:

> Life is full of choices between alternatives that offer different mixes of reinforcement and punishment. Going to work entails both getting paid (positive reinforcement) and suffering hassles (positive punishment), whereas calling in sick may forfeit some pay (negative punishment), avoid the hassles (negative reinforcement), allow a vacation (positive reinforcement), and incur some workplace disapproval (positive punishment). Which set of relations wins out depends on which relations are strong enough to dominate, and that depends on both the present circumstances and the person's history of reinforcement and punishment. (p. 68)

Vollmer (2002) suggested that the scientific study of punishment continue because unplanned and planned punishment occur frequently and that planned, sophisticated applications of punishment are within the scope of inquiry for applied behavior analysts. We believe that studying punishment is within the realm of a science of behavior. As Vollmer states, "Scientists and practitioners are obligated to understand the nature of punishment if for no other reason than because punishment happens." (p. 469)

Focused Research on Punishment Is Needed

Many inappropriate and ineffective applications of punishment may reflect our incomplete knowledge of this principle (Lerman & Vorndran, 2002). Our technical understanding about, and recommendations for, applying punishment are derived from

decades-old basic research. Although sound scientific data that describe still-relevant questions have an unlimited shelf life, additional basic research on the mechanisms and variables that produce effective punishment no doubt will uncover elements that remain tangled in complex variables and their interactions.

Basic laboratory research allows for the control of variables that is difficult or impossible to attain in applied settings. Once mechanisms are revealed in basic research, however, applications to real-world challenges may be devised or adapted with more confidence, but with the simultaneous recognition of limitations and ethics. Practitioners must acknowledge when, how, why, and under what conditions punishment produces behavioral suppression for the students, clients, or patients with whom they work. Further, we must come to the realization that while a laboratory finding, with all of its controls and high internal validity, might produce a resounding suppressive effect, an applied practitioner using an adaptation of that punishment application may only yield a fraction of that effect. Still, for the child who is self-mutilating, self-destructive, and seemingly only coming into contact with life-threatening situations, obtaining a fractional improvement in behavior suppression may be cause for celebrating progress.

There are at least two explanations for why it might be difficult to conduct experimental work on punishment. First, for a variety of reasons, Human Subjects Review Committees may be less than enthusiastic about supporting research on punishment. To the extent that fewer proposals to conduct basic research on punishment are approved, then fewer data will be available on its effects, side effects, and applicability in applied settings. Second, applied researchers, in response to complying with ethical considerations and the rights of individuals to least restrictive alternative treatments, have shifted their focus to positive reductive procedures or less intrusive negative punishment applications.

We support Horner's (2002) call for practical applications of punishment research in field settings. It is important to determine how punishment works in situations in which environmental variables may not be as well controlled and in which the practitioner delivering the punisher might not be a trained professional. Educational and clinical applications of punishment also demand that behavior analysts understand the individual, contextual, and environmental variables that produce effective application. Without a clear and refined knowledge base about these variables and conditions, applied behavior analysis as a science of behavior cannot make a legitimate claim to have achieved a comprehensive analysis of its own basic concepts (Lerman & Vorndran, 2002; Vollmer, 2002).

Interventions Featuring Unconditioned Punishers as Default Technologies

Iwata (1988) recommended that punishment-based interventions involving the contingent application of aversive stimulation be treated as default technologies. A default technology is one that a practitioner turns to when all other methods have failed. Iwata recommended that behavior analysts not advocate for the use of aversive technologies (because advocacy is not effective, not necessary, and not in the best interests of the field), but that they be involved in research and development of effective aversive technologies.

> We must do the work because, whether or not we like it, default technologies will evolve whenever there is failure and, in the case of aversive stimulation, we are in a unique position to make several contributions. First, we can modify the technology so that it is effective and safe. Second, we can improve it by incorporating contingencies of positive reinforcement. Third, we can regulate it so that application will proceed in a judicious and ethical manner. Last, and surely most important, by studying the conditions under which default technologies arise, as well as the technologies themselves, we might eventually do away with both. Can you think of a better fate for the field of applied behavior analysis? (p. 156)

SUMMARY

Definition and Characteristics of Punishment

1. Punishment has occurred when a response is followed immediately by a stimulus change that decreases the frequency of that type of behavior in the future.

2. Punishment is defined neither by the actions of the person delivering the consequences nor by the nature of those consequences. A decrease in the future occurrences of the behavior must be observed before a consequence-based intervention qualifies as punishment.

3. Positive punishment has occurred when the presentation of a stimulus (or an increase in the intensity of an already present stimulus) immediately following a behavior results in a decrease in the frequency of the behavior.

4. Negative punishment has occurred when the termination of an already present stimulus (or a decrease in the intensity of an already present stimulus) immediately following a behavior results in a decrease in future occurrences of the behavior.

5. Because aversive events are associated with positive punishment and with negative reinforcement, the term *aversive control* is often used to describe interventions involving either or both of these two principles.

6. A punisher is a stimulus change that immediately follows the occurrence of a behavior and reduces the future frequency of that type of behavior.

7. An unconditioned punisher is a stimulus whose presentation functions as punishment without having been paired with any other punishers.

8. A conditioned punisher is a stimulus that has acquired its punishing capabilities by being paired with unconditioned or conditioned punishers.

9. A generalized conditioned punisher will function as punishment under a wide range of motivating operations because of its previous pairing with numerous unconditioned and conditioned punishers.

10. A discriminative stimulus for punishment, or S^{Dp}, is a stimulus condition in the presence of which a response class occurs at a lower frequency than it does in the absence of the S^{Dp} as a result of a conditioning history in which responses in the presence of the S^{Dp} have been punished and similar responses in the absence of that stimulus have not been punished (or have resulted in a reduced frequency or magnitude of punishment).

11. Punishment's suppressive effects on behavior are usually not permanent. Recovery from punishment is especially likely when the behavior being punished also obtains reinforcement.

Factors That Influence the Effectiveness of Punishment

12. Results of basic and applied research show that punishment is more effective when
    - the onset of the punisher occurs as soon as possible after the occurrence of a target response,
    - the intensity of the punisher is high,
    - each occurrence of the behavior is followed by the punishing consequence,
    - reinforcement for the target behavior is reduced, and
    - reinforcement is available for alternative behaviors.

Possible Side Effects and Problems with Punishment

13. Punishment sometimes causes undesirable side effects and problems, such as the following:
    - Emotional and aggressive reactions
    - Escape and avoidance behaviors
    - Behavioral contrast: Reduced responding from punishment in one setting or situation may be accompanied by increased responding in another setting or situation in which responses go unpunished
    - Modeling of undesirable behavior
    - Overuse of punishment caused by the negative reinforcement of the punishing agent's behavior

Positive Punishment Interventions

14. Reprimands: Used sparingly, a firm reprimand such as "No!" can suppress future responding.

15. Response blocking: When the learner begins to emit the problem behavior, the therapist physically intervenes to prevent or "block" the completion of the response.

16. Response interruption and redirection (RIRD): This approach is a variation of response blocking that consists of interrupting stereotypic behavior at its onset and redirecting the individual to complete high-probability behaviors instead.

17. Contingent exercise: In this intervention, the client is required to perform a response that is not topographically related to the problem behavior.

18. Overcorrection: This is a punishment-based tactic in which, contingent on each occurrence of the problem behavior, the learner is required to engage in effortful behavior that is directly or logically related to the problem.

19. Restitutional overcorrection: In this intervention, the learner is required to repair the damage caused by the problem behavior and then to engage in additional behavior that brings the environment to a condition better than it was prior to the misbehavior.

20. Positive practice overcorrection: In this type of overcorrection, the learner repeatedly performs a correct form of the behavior, or a behavior incompatible with the problem behavior, for a specified time or number of responses.

21. Contingent exercise: In this intervention, the client is required to perform a response that is not topographically related to the problem behavior.

Guidelines for Using Punishment

22. To apply punishment with optimal effectiveness while minimizing undesirable side effects, a practitioner should:
    - Select effective and appropriate punishers: (a) conduct punisher assessments to identify the least intrusive punisher that can be applied consistently and safely; (b) use punishers of sufficient quality and intensity; (c) use a variety of punishers to combat habituation and increase the effectiveness of less intrusive punishers.
    - If problem behavior consists of a response chain, deliver the punisher as early in the response sequence as possible.
    - Punish each occurrence of the behavior.
    - If maintaining a continuous schedule is impractical, gradually shift to an intermittent schedule of punishment combined with extinction.
    - Mediate unavoidable response-to-punishment delays.
    - Supplement punishment with complementary interventions, in particular, differential reinforcement, extinction, and antecedent interventions.
    - Watch and be prepared for negative side effects.
    - Record, graph, and evaluate data daily.

Ethical Considerations Regarding the Use of Punishment

23. The first ethical responsibility for any human services professional or agency is to do no harm. Any intervention

must be physically safe for all involved and contain no elements that are degrading or disrespectful to the client.

24. The doctrine of the least restrictive alternative holds that less intrusive procedures (e.g., positive reductive approaches) must be tried first and found to be ineffective before more intrusive procedures are implemented (e.g., a punishment-based intervention).

25. A client's right to effective treatment raises an important ethical issue. Some maintain that the failure to use a punishment procedure that research has shown to suppress self-destructive behavior is unethical because it withholds a potentially effective treatment and may maintain a dangerous or uncomfortable state for the client.

26. Agencies and individuals providing applied behavior analysis services can help ensure that applications of punishment-based interventions are safe, humane, ethical, and effective by creating and following a set of policy standards, procedural safeguards, and evaluation requirements.

Concluding Perspectives

27. Applied behavior analysts should recognize and appreciate the natural role of punishment and the importance of punishment to learning.

28. Many misapplications of punishment reflect the field's incomplete knowledge of the principle. More basic and applied research on punishment is needed.

29. Applied behavior analysts should advocate for the judicious and ethical application punishment, incorporate a full range of differential reinforcement contingencies, antecedent interventions, and extinction to reduce behaviors, study the conditions under which default technologies are considered and used, so as to eventually make them unnecessary.

KEY TERMS

behavioral contrast
conditioned punisher
discriminative stimulus for punishment
generalized conditioned punisher
negative punishment
overcorrection
positive practice overcorrection
positive punishment
punisher
punishment
recovery from punishment
response blocking
response interruption and redirection (RIRD)
restitutional overcorrection
unconditioned punisher

MULTIPLE-CHOICE QUESTIONS

1. Which of the following statements describes the definition of positive punishment?
   a. The presentation of a stimulus that increases the rate of occurrence of the target behavior
   b. The presentation of a stimulus that decreases the rate of occurrence of the target behavior
   c. The removal of a stimulus that decreases the rate of occurrence of the target behavior
   d. The removal of a stimulus that increases the rate of occurrence of the target behavior

   Hint: (See "Positive Punishment and Negative Punishment")

2. Which of the following statements describes the definition of negative punishment?
   a. The presentation of a stimulus that increases the rate of occurrence of the target behavior
   b. The presentation of a stimulus that decreases the rate of occurrence of the target behavior
   c. The removal of a stimulus that decreases the rate of occurrence of the target behavior
   d. The removal of a stimulus that increases the rate of occurrence of the target behavior

   Hint: (See "Positive Punishment and Negative Punishment")

3. Which statement best describes recovery from punishment?
   a. After punishment is discontinued, sometimes the behavior that experienced the punishment contingency will recover.
   b. When the punishment contingency is implemented, the behavior of interest may initially increase in rate, followed by a rapid decline in frequency.
   c. After punishment is discontinued, the behavior of interest may gradually reappear, only to disappear on its own, regardless of existing environmental contingencies.
   d. The period of time the implementer of the punishment contingency requires to "recover" from the emotional side effects of using punishment.

   Hint: (See "Recovery from Punishment")

4. Which of the following is not a factor that influences the effectiveness of punishment listed in the chapter?
   a. Immediacy
   b. Intensity
   c. Variation
   d. Schedule

   Hint: (See "Factors That Influence the Effectiveness of Punishment")

5. The use of a punishment procedure may cause unwanted side effects to appear. The chapter presents several side effects associated with the use of punishment. Which of the following is not a side effect listed in the chapter?
   a. Undesirable emotional responses
   b. Avoidance and escape
   c. Increased rate of problem behavior
   d. Decreased self-esteem
   Hint: (See "Possible Side Effects and Problems with Punishment")
6. Little Peter was walking along the fence of his uncle's dairy farm. Peter reached up and grabbed the wire along the fence, not realizing it was electrified. After recovering from the pain caused by the shock, Peter never again touched the fence wires. Which statement best describes the shock stimulus felt by Peter?
   a. Negative Reinforcement
   b. Negative Punishment
   c. Bad Judgment
   d. Positive Punishment
   Hint: ("Positive Punishment Interventions")
7. Which statement best describes positive punishment?
   a. The delivery of a stimulus, following a behavior, which increases the likelihood that the behavior will occur again
   b. The delivery of a stimulus, following a behavior, which decreases the likelihood that the behavior will occur again
   c. A stimulus delivered with good intentions
   d. The removal of a stimulus, following a behavior, which decreases the likelihood that the behavior will occur again
   Hint: (See "Positive Punishment Interventions")
8. Which statement best describes the procedure of response blocking?
   a. Blocking the occurrence of the antecedent stimulus associated with the occurrence of the problem behavior
   b. Delivering a wood block to the student prior to the occurrence of a behavioral response
   c. Physically intervening as soon as the student begins to emit a problem behavior to prevent the completion of the response
   d. The interventionist closing his or her eyes to deny the student pleasure after engaging in the problem behavior
   Hint: (See "Response Blocking")
9. Punishment should be thought of as an "eye-for-an-eye" procedure.
   a. True
   b. False
   Hint: (See "Guidelines for Using Punishment")
10. Which statement is not recommended as a guideline for the use of punishment?
   a. Test punisher on your pet prior to implementing it with your student
   b. Conduct punisher assessment
   c. Punish each occurrence of the behavior
   d. Deliver the punisher immediately
   Hint: (See "Guidelines for Using Punishment")
11. Carr and Lovaas (1983) recommended that practitioners experience any punisher personally before the treatment begins for what reason?
   a. It reminds the practitioner that it is a sadistic procedure.
   b. It reminds the practitioner that the technique produces physical discomfort.
   c. It reminds the practitioner that the technique is designed to increase the appropriate behavior of the student.
   d. It allows the practitioner to determine a baseline of discomfort which he/she then increases rapidly as it is administered to the student.
   Hint: (See "Guidelines for Using Punishment")
12. Which of the following is not listed in the chapter as an ethical consideration regarding the use of punishment?
   a. Right to safe and humane treatment
   b. Least restrictive alternative
   c. Right to effective treatment
   d. Most restrictive alternative
   Hint: (See "Ethical Considerations Regarding the Use of Punishment")
13. Sean was responsible for designing an intervention for a fifth-grade student who engaged in severe problem behavior consisting of throwing chairs at the teacher when an assignment was given. Sean decided to utilize a contingent physical restraint every time the child threw a chair. Sean is violating the ethical consideration of:
   a. Right to safe and humane treatment
   b. Least restrictive alternative
   c. Right to effective treatment
   d. Sean is not violating any ethical considerations
   Hint: (See "Ethical Considerations Regarding the Use of Punishment")

14. Which statement regarding the strength of applied behavior analysis is not supported in the punishment chapter?

a. Recognize and appreciate punishment's natural role and contributions to survival and learning

b. Increase the use of punishment procedures over the use of reinforcement-based interventions

c. Increase basic and applied research on punishment

d. View treatments featuring positive punishment as default technologies

Hint: (See "Concluding Perspectives")

15. Many of the recommendations for punishment are derived from basic research conducted more than _____ years ago.

a. 5

b. 20

c. 10

d. 40

Hint: (See "Concluding Perspectives")

ESSAY-TYPE QUESTIONS

1. Explain the procedures of positive punishment and negative punishment. Be sure to include in your answer factors that may influence the effectiveness of the punisher.

   Hint: (See "Definition and Characteristics of Punishment")

2. Discuss the possible side effects of using punishment.

   Hint: (See "Possible Side Effects and Problems with Punishment")

3. Define and give examples of the five types of positive punishment given in the chapter: reprimands, response blocking, contingent exercise, overcorrection, and response interruption and redirection.

   Hint: (See "Positive Punishment Interventions")

4. With the decision to use punishment-based treatment, everything should be done to ensure that the punishment is as effective as possible. Identify and discuss the suggested guidelines for the use of punishment.

   Hint: (See "Guidelines for Using Punishment")

5. The use of punishment presents some major ethical concerns. Identify and discuss two of the ethical concerns associated with the use of punishment and give an example of a situation in which the concern is involved.

   Hint: (See "Ethical Considerations Regarding the Use of Punishment")

6. Most of the punishment research is derived from basic studies that occurred more than 40 years ago. Discuss and contribute to Horner's (2002) call for practical applications of punishment research in field settings.

   Hint: (See "Concluding Perspectives")

NOTES

1. Although the term *automatic punishment* is used infrequently in the behavior analysis literature, it is similar to *automatic reinforcement*. Automatic punishment occurs when a punishing consequence (e.g., burned finger) is a socially unmediated, unavoidable outcome of a response (e.g., touching a hot stove).
2. A stimulus that becomes a conditioned punisher does not have to be a neutral stimulus prior to its pairing with a punisher. The stimulus could already function as a reinforcer under other conditions. For example, the onset of a blue light that has been paired repeatedly with reinforcement in one setting and with punishment in another setting is a conditioned reinforcer or a conditioned punisher, depending on the setting.
3. Instructions or conditioned stimuli that have been paired with known aversives can mitigate the time-delay gap.
4. Multiple schedules of reinforcement are described in Chapter 13.
5. When an extinction procedure is first implemented, an increase in responding, called an *extinction burst*, is sometimes observed before the response rate begins to decline. The principle, procedure, and effects of extinction are detailed in Chapter 22.
6. Increasing the effort or force required to perform a behavior can also be an effective tactic for reducing responding (Friman & Poling, 1995). There is no consensus as to whether punishment accounts for the reduced responding. As with response blocking, one perspective from which increased response effort can be conceptualized as a punishment procedure is to consider the movement necessary to come into contact with the increased effort requirement as a member of the target behavior response class. In that case, the increased effort required to continue the response to completion is (a) a consequence for the response that brought the learner into contact with it and (b) aversive stimulation that functions as punishment as the frequency of future responding decreases.
7. Chapter 31 discusses ethical issues and practices for applied behavior analysts.

CHAPTER 15

Negative Punishment

LEARNING OBJECTIVES

- Define negative punishment.
- Define nonexclusion time-out.
- Define exclusion time-out.
- State the procedures for implementing planned ignoring.
- State the procedures for implementing contingent observation.
- State the procedures for implementing withdrawal of a specific positive reinforcer.
- State the procedures for using a time-out room.
- State the procedures for implementing a partition time-out.
- List the decisions a practitioner must make prior to, during, and after a time-out application.
- Define response cost.
- List the undesirable aspects of negative punishment.

Half a minute into watching the funniest Internet video ever, Ruth clicked a button. Instead of seeing the full-screen view that she expected, the video disappeared. She reopened the video, advanced the cursor to the previous time stamp, clicked the same button, and once again lost access to the video. Ruth never again clicked the button that makes videos disappear!

Kent selected the last two front-row tickets for an upcoming concert by his favorite band. After entering his credit card information and billing address, he was prompted to review the details before clicking the purchase button. Just then Kent's cell phone buzzed with a text message, which led to several minutes of exchanged messages. When Kent returned to the ticket seller's website, the two tickets were no longer available. Kent no longer accepts cell phone texts when completing online purchases.

Negative punishment contingencies changed the behavior of both actors. Ruth no longer clicks the button that causes videos to disappear; Kent does not reply to text messages when shopping online. In **negative punishment**, a response is followed immediately by the *removal* of a stimulus that decreases future occurrences of similar responses. By contrast, in positive punishment, a response is followed immediately by the *presentation* of a stimulus that decreases future occurrences of similar responses. Both negative and positive punishment contingencies decrease future occurrences of behaviors. The distinction is in the type of stimulus change—removal or presentation—that functions as the punishing consequence (see Figures 15.1 and 14.2).

Applied behavior analysts implement negative punishment with two principal tactics: time-out from positive reinforcement and response cost. This chapter defines and operationalizes time-out and response cost, describes how these negative punishment tactics are used in applied settings, outlines desirable and undesirable aspects of each procedure, and offers practical and ethical guidelines to help practitioners implement time-out and response cost effectively.

TIME-OUT FROM POSITIVE REINFORCEMENT DEFINED

Time-out from positive reinforcement (or simply **time-out**) is the immediate response-contingent withdrawal of the opportunity to earn positive reinforcers or the immediate loss of access to positive reinforcers for a specified time. To function as negative punishment, the time-out procedure must decrease future occurrences of similar behaviors. Implicit in the definition of time-out are three important factors: (a) The discrepancy between the "time-in" and the time-out environments must be discriminative, (b) the response-contingent loss of access to reinforcement must occur immediately, and (c) a resultant decrease in the future frequency of the time-out–producing behavior must occur.[1]

Time-out can be viewed from a procedural, conceptual, or functional perspective. Procedurally, time-out entails removing the person from a reinforcing environment, or prohibiting access to reinforcers within his current environment for a specified period of time contingent upon the occurrence of problem behavior.

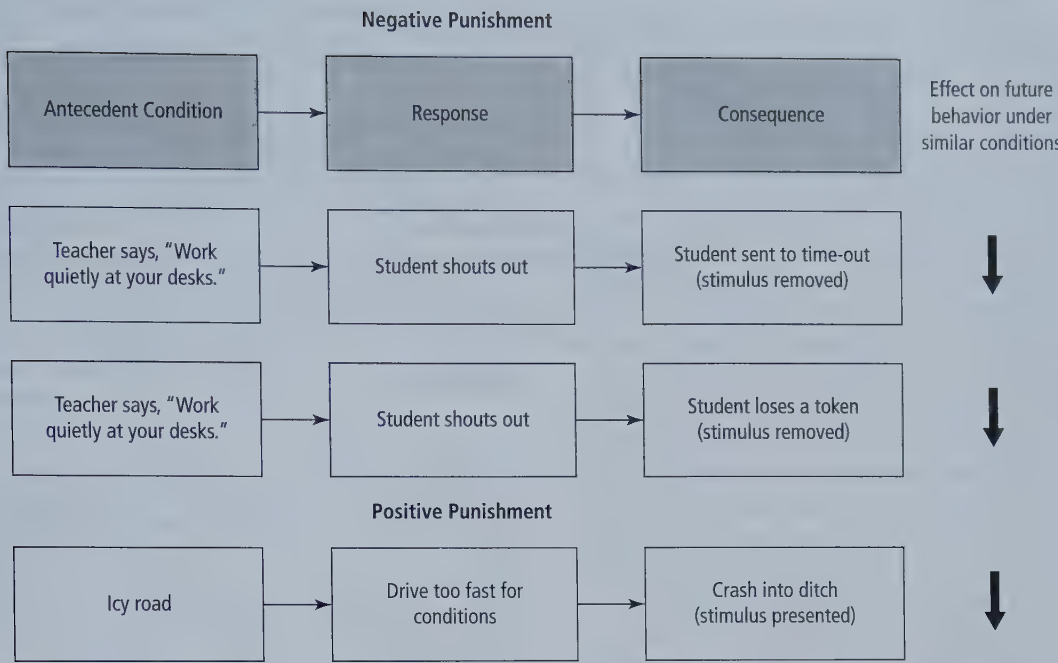


Figure 15.1 Negative punishment and positive punishment are distinguished by the type of stimulus change that functions as the punishing consequence.

Conceptually, the client must discriminate between the time-out and the time-in environment: that is, the more discriminatively reinforcing the time-in setting, the more effective time-out is as a negative punisher (Solnick, Rincover, & Peterson, 1977). In other words, the greater the difference between the reinforcing value of time-in and absence of that reinforcing value in the time-out setting, the more effective time-out will be.

From a functional perspective, time-out involves reduced future occurrences of similar behaviors. Without a reduction of future similar behavior, time-out is not in effect, even if the person has been removed from the setting, segregated briefly from the group but remains in the setting, or loses access to reinforcers. Assuming an exclusion time-out procedure was intended, if a teacher removes a student from the classroom—presumably the time-in setting—but upon the student's return, the problem behavior continues, effective time-out cannot be claimed.

So, contrary to popular thinking, time-out does not involve simply removing an individual to a sequestered setting. Although such removal may arguably describe an element of isolation, exclusion, or seclusion (Costenbader & Reading-Brown, 1995), the mere removal to another location falls far short of being accurately named time-out from positive reinforcement. Procedural, conceptual, and functional criteria need to be met to assert that claim.

TIME-OUT TACTICS FOR APPLIED SETTINGS

There are two basic categories of time-out available in applied settings: nonexclusion and exclusion. Within each type, several procedural variations provide options for the practitioner when deciding on a course of action to reduce similar behavior in the future (see Figure 15.2). As a rule, nonexclusion time-out should be the method of first choice because practitioners are ethically bound to employ the most powerful, but least restrictive, alternative when deciding on eligible variations. Regardless of which category is employed, we believe that, as a rule, time-out sessions should be brief, lasting not more than 2 to 5 minutes.

Nonexclusion Time-Out

Nonexclusion time-out means that contingent on the occurrence of the target behavior, the participant remains physically within the time-in setting, but elements of that setting are changed so that (a) the opportunity to earn reinforcers is removed; (b) a specific reinforcer is terminated; (c) repositioning occurs so that observations of ongoing activity continue, but without reinforcement; or (d) the individual is removed to a different space within the time-in setting.

Nonexclusion time-out is typically implemented using one of four tactics: planned ignoring, terminate specific reinforcer contact, contingent observation, and partition/select space time-out.

Planned Ignoring

Planned ignoring occurs when, contingent on the occurrence of an inappropriate behavior, the opportunity to earn social reinforcers—usually attention, physical contact, or verbal interaction—is removed. Operationally, planned ignoring involves systematically looking away from the individual, remaining quiet, or refraining from any interaction whatsoever for a specific time (Kee, Hill, & Weist, 1999). For example, if a behavior targeted for time-out occurred in a one-to-one instructional session, the therapist would simply turn her back to the learner for the duration of the time-out interval. The therapist's turning away essentially removes the time-in environment and denies the learner access to reinforcers.

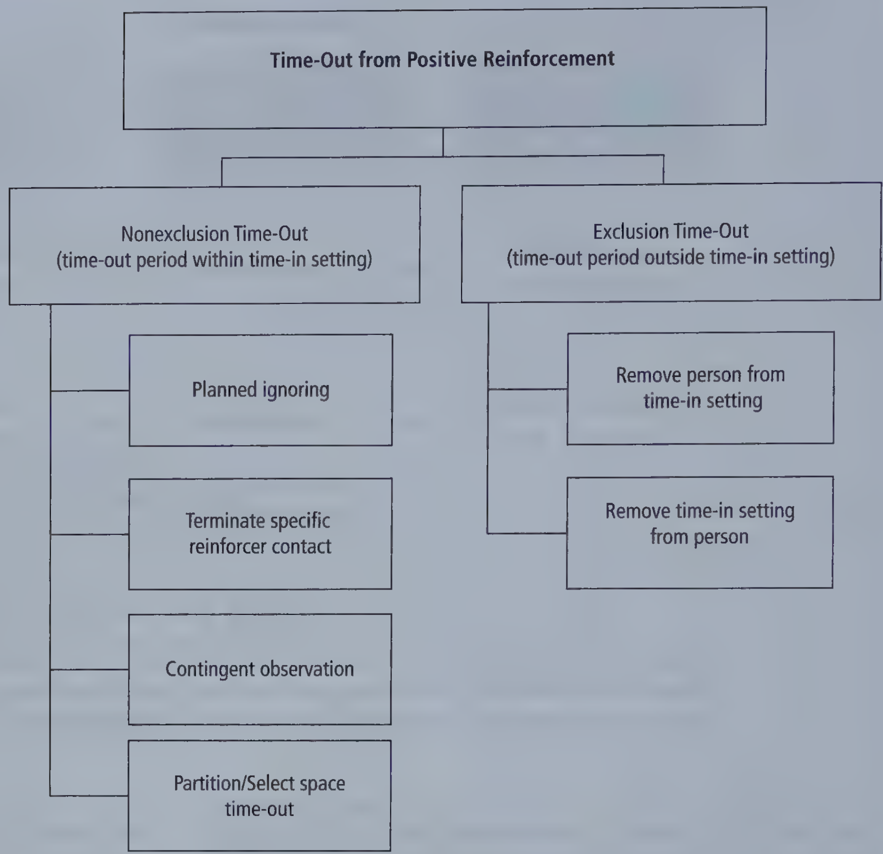


Figure 15.2 Tactics for implementing time-out from positive reinforcement.

Terminate Specific Reinforcer Contact

Terminate specific reinforcer contact is a variation of time-out whereby each occurrence of the target behavior immediately stops an activity or sensory reinforcer. West and Smith (2002) reported an application of this tactic to reduce noise level during school lunchtime. They mounted a facsimile traffic signal light to the cafeteria wall and fitted the light with a sensor to detect noise levels. When the sensor detected an acceptable conversation decibel level, a green light registered on the traffic signal and music played through cafeteria speakers (i.e., music was identified as a reinforcer and constituted the time-in condition). The music played as long as the students engaged in an appropriate level of conversation. If the conversational decibel level increased to a predetermined threshold, the traffic light changed from green to yellow, visibly warning the students that further increases in noise would lead to time-out from music. If the decibel level rose still higher, the traffic signal turned from yellow to red, and the music ceased for 10 sec. Under this group procedure, louder decibel conversations were reduced.

Bishop and Stumphauzer (1973) demonstrated that the contingent termination of television cartoons successfully reduced thumb sucking in three young children. Unknown to the children, a remote on/off switch was attached to the television. Baseline data indicated that each child emitted a high rate of thumb sucking while viewing the cartoons. During the terminate contact with a specific reinforcer variation of time-out, the television was immediately turned off when thumb sucking occurred, and it was turned back on when thumb sucking stopped. The procedure was effective in reducing thumb sucking not only in the treatment location (an office) but also during a story period at school.

Contingent Observation

In **contingent observation**, the person is repositioned within an existing setting such that he or she can observe ongoing activities, but access to available reinforcers is lost. A teacher uses contingent observation after an undesirable behavior when she redirects the offending student to move from the group, but remain visible, and reinforcers are withheld for a specified time. In short, the student is directed to "sit and watch" (Twyman, Johnson, Buie, & Nelson, 1994; White & Bailey, 1990). When the contingent observation period ends, the student rejoins the group, and is then able to earn reinforcers for appropriate behavior. An athletic analog to contingent observation occurs when an athlete commits an unnecessary foul during a game, and, in turn, the coach immediately "sidelines" the player. The player can still see the game; however, he cannot participate in it, or receive social reinforcers for contributing to the outcome (e.g., crowd cheers, praise from fellow players).

Figure 15.3 shows the effects of White and Bailey's (1990) use of contingent observation in reducing the number of disruptive behaviors in two inclusive physical education classes. When contingent observation was in effect (i.e., sit and watch), the number of disruptive behaviors across two classrooms decreased and remained near zero levels.

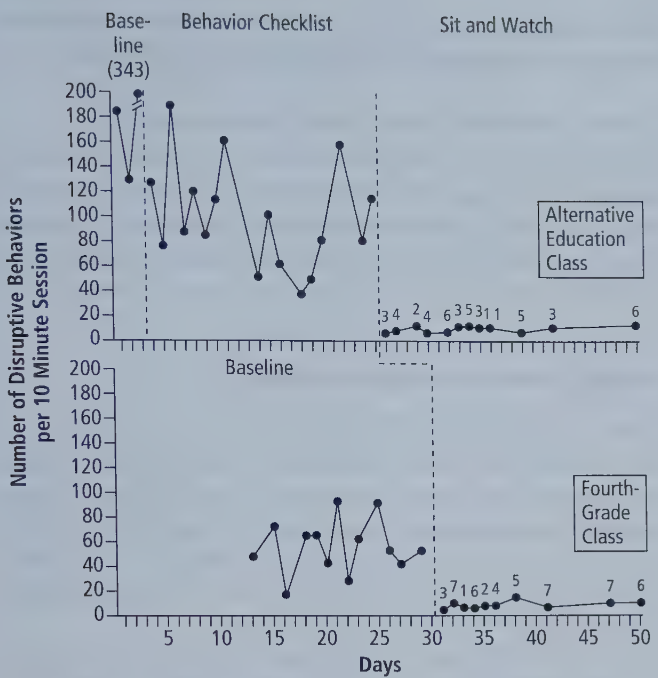


Figure 15.3 Number of disruptive behaviors per 10-min observation period. The numbers above the data points represent the number of times "sit and watch" was implemented during the class.

From "Reducing Disruptive Behaviors of Elementary Physical Education Students with Sit and Watch", A. G. White and J. S. Bailey, 1990, *Journal of Applied Behavior Analysis, 23*, p. 357. Reproduced with permission of John Wiley & Sons Inc.

Partition/Select Space Time-Out

In **partition or select space time-out**, the student remains within the time-in setting, but his view within the setting is restricted by a panel or cubicle, or a select space is arranged to serve as the time-out area (i.e., a carpet, a corner). A teacher would use partition time-out by directing a student, contingent on a misbehavior, to move from his assigned seat to that partitioned location for a specified time period. Although partition or select space time-out has the advantage of keeping the student within the time-in setting—presumably to hear academic content and the teacher praising other students for appropriate behavior—it also can be disadvantageous. That is, the individual may still be able to obtain reinforcement from other students, reducing effectiveness. Second, the student might become the object of disparaging comments from other students (e.g., "Brandon is a dummy; he's in trouble."). Third, since the individual's view—and the teacher's view—can be at least partially restricted, undesirable surreptitious behaviors may occur. Also, public perceptions must again be taken into account with this form of exclusion. Finally, even though the student remains in the room, the partition time-out area might be viewed as exclusionary.

Exclusion Time-Out

Exclusion time-out is the physical separation of the participant from the time-in environment when time-out goes into effect. There are two basic tactics for implementing exclusion time-out contingent on the occurrence of the target behavior: (a) The participant is removed immediately from the time-in environment, or (b) the time-in environment is removed immediately from the participant.

Participant Removed from Time-In Setting

A *time-out room* is any confined, safe, secure, supervised, and continuously monitored area within the participant's normal educational or treatment facility that is devoid of existing positive reinforcers, that does not have access to potential positive reinforcers, and in which the person can be securely placed. A time-out room should preferably be located in close proximity to the time-in setting, and it should have minimal furnishing (e.g., a chair and a table). It should have adequate light, heat, safety features, and ventilation, but should not contain potentially reinforcing items (e.g., pictures on the wall, telephones, breakable objects, windows with access to playgrounds, or busy streets). The room should not be locked, and it must be monitored continuously by staff.

A time-out room has several advantages that make it attractive to practitioners. First, the opportunity to acquire reinforcement during time-out is eliminated or reduced substantially because the time-out environment is physically constructed to minimize such stimuli. Second, after a few exposures to the time-out room, students learn to discriminate this room from other rooms in the building. The room assumes conditioned aversive properties, thus increasing the probability that the time-in setting will be more reinforcing. Third, the risk of a student hurting other students in the time-in setting is reduced when the offending student is removed to this space.

However, before using a time-out room, the practitioner must consider several strong and compelling disadvantages. Escorting a student to the time-out room is not without its potential challenges: Practitioners can encounter resistance and outbursts, and they should also anticipate and be fully prepared to deal with emotional escalations. Using a time-out room may serve as a negative reinforcer for the teacher, thus increasing its use (Skiba & Raison, 1990). Also, the time-out room may set the occasion for the person to engage in undetected self-destructive or self-stimulatory behaviors that should be stopped (Solnick et al., 1977). This situation compromises the effectiveness of time-out. Practitioners must be sensitive to the public's perception of the time-out room. Even a benign-appearing time-out room can be viewed with distrust by persons who perceive the intent of the room differently and are misinformed about its purpose.

In recent years, the use of the time-out room has increasingly fallen into disfavor with practitioners, members of professional associations, and client rights advocacy groups, all of whom embrace less intrusive, but still effective, interventions for problem behavior.[2]

Hallway time-out used to be common practice in schools, and some teachers continue to employ it, despite increasing administrative mandates to prohibit it. In *hallway time-out*, the student is directed to leave the classroom and sit (or stand) outside the classroom in the hallway or corridor until the time-out period expires. We do not recommend this procedure for several reasons: (a) Many school districts and child service agencies prohibit teachers and staff from placing children in

unsupervised settings; (b) the student can obtain "bootleg" reinforcement from a multitude of sources (e.g., students in other rooms, individuals walking in the hallway, school visitors); (c) there is increased likelihood of the student escaping (i.e., the student may run away from the school altogether); (d) the teacher is less likely to be able to monitor the student outside of her line of sight; and (e) unattended students, especially those in the midst of a behavioral episode, may represent a safety hazard to themselves or others. Therefore, for ethical, safety, and programmatic reasons, we strongly recommend that practitioners seek alternatives to hallway time-out when trying to decrease problematic behavior.

Time-In Setting Removed from Participant

This variation of exclusion time-out does not require changing the participant's physical location. Instead, contingent on the occurrence of the target behavior, the time-in environment is removed from the participant. This form of time-out is a feature of some computer-based learning programs whereby certain error responses are followed immediately by the screen going dark (or blank) for several seconds, denying the user access to the activity, and, hence, its reinforcers.

Desirable Aspects of Time-Out

Ease of Application

Time-out, especially in its nonexclusion variations, is relatively easy to apply. The teacher can accomplish even physically removing a student from the environment with comparative ease if the teacher acts in a businesslike fashion and does not attempt to embarrass the student. Issuing a direction privately (e.g., "Deion, you have disrupted Monique twice; time-out is now in effect.") can help the teacher handle a student who has misbehaved but who does not want to leave the room. If the behavior warrants exclusion time-out, the teacher must insist that the student leave; however, the teacher should communicate this insistence at close range so that the student is not placed in a position of challenging the teacher openly to save face with a peer group. Teachers should consult district policy on whether an administrator, or other staff member, should be called to assist with escorting the student to the time-out location.

Acceptability

Time-out, especially in its nonexclusion variations, meets an acceptability standard because practitioners regard it as appropriate, fair, and effective. Even so, practitioners should always check with the appropriate administering body to ensure compliance with school or agency policy before applying time-out for major or minor infractions.

Rapid Suppression of Behavior

When effectively implemented, time-out suppresses the target behavior in a moderate to rapid fashion. Sometimes only a few applications are needed to achieve acceptable reduction levels. Other reductive procedures (e.g., extinction, differential reinforcement of low rates) also produce decreases in behavior, but they can be time-consuming. Practitioners may not have the luxury of waiting several days or a week for a behavior to decrease. In such instances, time-out merits consideration.

Can Be Combined with Other Applications

Time-out can be combined with other procedures, extending its usability in applied settings (Dupuis, Lerman, Tsami, & Shireman, 2015). When combined with differential reinforcement, time-out can increase desirable behavior and decrease undesirable behavior (Byrd, Richards, Hove, & Friman, 2002).

USING TIME-OUT EFFECTIVELY

Although time-out has shown effectiveness in reducing behavior, it should not be the method of first choice. Practitioners faced with the task of reducing behavior should initially consider nonpunitive methods for reducing the target behavior (e.g., differential reinforcement of alternative behavior, extinction). Time-out should be considered only when the evidence shows that less intrusive procedures have failed.

Further, effective implementation of time-out requires the practitioner to make several decisions prior to, during, and after time-out implementation. Figure 15.4 shows Powell and Powell's (1982) time-out implementation checklist as an aid in the decision-making process. The following sections expand on several major points underlying many of these decisions.

Enrich the Time-In Environment

Selected behaviors in the time-in environment must be reinforced. In choosing the behaviors to reinforce during the time-in setting, practitioners should identify behaviors that are alternative to, or incompatible with, behaviors that will lead to time-out (e.g., using differential reinforcement of alternative behavior, differential reinforcement of incompatible behavior). Differential reinforcement will facilitate the development of appropriate behaviors. Further, the ratio of reinforcing statements to reprimanding comments should strongly fall in favor of reinforcing statements. Additionally, upon return from time-out, reinforcement for appropriate behavior should be delivered as quickly as possible.

Assuming that the effectiveness of any reinforcing or punishing stimulus is derived from the antecedent-behavior-consequence contingency, practitioners should examine all aspects of this relationship, including the time-in setting. Answering the following questions will help in that regard.

- Are the presumed reinforcers actually functioning as reinforcers?
- What function does the target behavior serve (i.e., attention, escape)?
- Is the time-out setting appropriate, and is the contingency being delivered immediately?

Within this context, practitioners are then charged with individualizing time-out to the specific circumstances at hand. Solnick et al. (1977) provide guidance: "This view of timeout suggests that there can be no 'standard' timeout procedure that will reliably reduce problem behavior. Alternatively, we are left

Figure 15.4 Steps for planning, implementing, evaluating, and terminating time-out.

Step	*Task*	*Date Completed*	*Teacher's Initials*
1.	Try less aversive techniques and document results.	_____	_____
2.	Operationally define disruptive behaviors.	_____	_____
3.	Record baseline on target behaviors.	_____	_____
4.	Consider present levels of reinforcement (strengthen if necessary).	_____	_____
5.	Decide on time-out procedure to be used.	_____	_____
6.	Decide on time-out area.	_____	_____
7.	Decide on length of time-out.	_____	_____
8.	Decide on command to be used to place child in time-out.	_____	_____
9.	Set specific criteria that will signal discontinuance of time-out.	_____	_____
10.	Set dates for formal review of the time-out procedure.	_____	_____
11.	Specify back-up procedures for typical time-out problems.	_____	_____
12.	Write up the entire procedure.	_____	_____
13.	Have the procedure reviewed by peers and supervisors.	_____	_____
14.	Secure parental/guardian approval and include the written program in the child's IEP.	_____	_____
15.	Explain procedure to the student and class (if appropriate).	_____	_____
16.	Implement the procedure, take data, and review progress daily.	_____	_____
17.	Formally review procedure as indicated.	_____	_____
18.	Modify the procedure as needed.	_____	_____
19.	Record results for future teachers/programs.	_____	_____

From "The Use and Abuse of Using the Timeout Procedure for Disruptive Pupils" by T. H. Powell and I. Q. Powell, 1982, *The Pointer*, *26*(2), p. 21. Reprinted by permission of the Helen Dwight Reid Educational Foundation. Published by Heldref Publications, 1319 Eighteenth St., N.W., Washington DC 20036-1802. Copyright © 1982.

with the task of monitoring the relative reinforcing properties of the timein and timeout setting for each child" (p. 423).

Define the Behavior That Will Lead to Time-Out

Before time-out is implemented, all of the appropriate parties must be informed of the target behaviors that will lead to time-out. If a teacher decides to use time-out, she should describe in explicit, observable terms the behavior that will result in time-out. Merely informing students that disruptive behavior will result in time-out is not sufficient. Giving specific examples and nonexamples of what is meant by disruptive behavior is a better course to follow. Donaldson and Vollmer (2011) provided operational definitions using measurable criteria for screaming, disrupting, and aggressive behaviors. It is important to measure behaviors precisely because in post–time-out interviews, Readdick and Chapman (2000) discovered that children were not always aware of the reasons time-out was implemented. Providing explicit examples and nonexamples addresses this problem.

Determine the Form and Variation of Time-Out

School board policy or institutional constraints may be established that set the parameters for time-out variations that can be used in applied settings. In addition, physical factors within the building (e.g., a lack of available space) may prohibit exclusionary forms of time-out. As stated previously, the behavior analyst should first consider nonexclusion time-out tactics.

Obtain Permission

One of the more important tasks a practitioner must perform prior to using time-out, especially in its exclusionary variations, is to obtain permission from supervisors or administrators. Because of the potential to misuse time-out (e.g., leaving the person in time-out too long, continuing to use time-out when it is not effective), practitioners must obtain administrative approval before employing it. However, interactions can happen so fast in applied settings that obtaining permission on an instance-by-instance basis would be unduly cumbersome. A preferred method is for the practitioner, in cooperation with administrators, to decide beforehand the type of time-out that will be implemented for certain offenses (e.g., nonexclusion versus exclusion), the duration of time-out (e.g., 2 to 5 minutes), and the behaviors that will demonstrate reinstatement to the time-in environment. Communicating these procedures and/or policies to other practitioners and caretakers in the building, as well as to parents, is advisable.

Explain the Time-Out Procedure and Rules

In addition to posting the behaviors that will lead to time-out, the teacher should also post the rules. Posting helps make explicit the behaviors that produce the contingencies (Van Houten, 1979). Minimally, these rules should focus on the initial duration of time-out and the exit criteria (i.e., rules that determine when time-out is in effect, when it is over, and what happens if ongoing inappropriate behavior occurs when the scheduled time-out period ends).

Determine the Duration of the Time-Out Period

Time-out periods should be as brief as possible while maintaining effectiveness. Time-out durations of 2 to 5 minutes are often effective in reducing behavior, although a short duration may be ineffective initially if an individual has had a history of longer time-out periods. As a rule, time-out periods exceeding 15 minutes are not effective.

Longer time-out periods are counterproductive for several reasons. First, the person may develop a tolerance for the longer duration and find ways to obtain reinforcers during time-out. Observing the person unobtrusively in time-out may provide evidence of the reinforcers being obtained (Solnick et al., 1977). This situation occurs with people who have a history of self-stimulatory behavior. The longer the duration of time-out, the more opportunity there is to engage in self-stimulation, and the less effective time-out becomes. Second, longer time-out periods remove the person from the educational, therapeutic, or family time-in environment in which the opportunity to learn and earn reinforcers is available. Third, given the undesirable practical, legal, and ethical aspects of longer time-out periods, a prudent course of action is for the practitioner to use relatively short, but consistent, time-out periods. When these periods are short, student academic achievement is not affected adversely (Skiba & Raison, 1990).

Apply Time-Out Consistently

Each occurrence of the undesirable behavior should lead to time-out. If a teacher, parent, or therapist is not in a position to deliver the time-out consequence after each occurrence of the target behavior, it may be better to use an alternative reductive technique. Using time-out occasionally may lead to student or client confusion about which behaviors are acceptable and which are not.

Establish Exit Criteria

Practitioners can consider exit criteria options to include: (a) extending the time-out period for a predetermined amount of additional time, (b) continuing the time-out period until the disruptive behavior ceases, or (c) requiring the absence of misbehavior for a specified time prior to the end of the fixed time-out period (e.g., no misbehavior for the last 30 seconds of the 5-min period). This latter tactic is termed a *release contingency*. The rationale for implementing a release contingency is twofold: (1) to mitigate the risk of adventitiously reinforcing the behavior that leads to time-out by releasing the person while that behavior is being emitted (i.e., an escape contingency); and (2) to reduce the effects of the time-out behavior from carrying over, or spilling into, the time-in setting. A release contingency has been evaluated in several studies (e.g., Donaldson & Vollmer, 2011; Erford, 1999; Mace, Page, Ivancic, & O'Brien, 1986).

Donaldson and Vollmer (2011) compared fixed-duration versus a release contingency time-out procedures. Participants were four preschool children who exhibited high rates of problem behavior (e.g., screaming, crying, aggression, disruptions). The researchers collected data on (a) behaviors that led to time-out and (b) delay-producing responses (i.e., problem behavior that occurred during the time-out period). During baseline, teachers, parents, and experimenters conducted regularly scheduled playground or circle time activities as usual. Children were expected to participate in a variety of low-demand activities, and programmed consequences were not delivered should a disruptive episode occur. If children engaged in potentially dangerous behavior—jumping from playground equipment, for example—teachers or experimenters intervened by blocking and redirecting, but not by reprimanding.

During the fixed-duration time-out condition, the child was shown a yellow card prior to the start of each session and told, "If you are sent to time-out, you only have to stay for 4 minutes, no matter what" (p. 698). The yellow card was then placed in a visible location, and each occurrence of problem behavior resulted in a fixed-duration 4-min time-out.

During the release contingency time-out condition, the participant was shown a red card prior to the start of each session and told, "If you are sent to time-out, you will have to stay for 4 minutes, but you cannot leave time-out until you are calm" (p. 698). The red card was placed where the child could see it, and each occurrence of problem behavior resulted in a fixed-duration 4-min time-out. However, if disruptive behavior occurred during the last 30 seconds of that 4-min interval, time-out was extended until 30 seconds elapsed without problem or until 10 minutes passed.

Both time-out procedures reduced problem behavior by all four participants (see Figure 15.5). Donaldson and Vollmer (2011) reported that problem behavior that occurred during either the fixed-time or the release contingency time-out periods did not predict spillover effects in the classroom, home, or playground after time-out periods were terminated. They concluded the discussion of the study's results with this recommendation: "A fixed duration time-out should be considered first because there was no apparent benefit to including a release contingency, release contingencies resulted in longer time-outs, and the effort required to implement a release contingency time-out is greater than that required to implement a fixed-duration time-out" (p. 704).

The decision to exit time-out should consider not only the child's behavior during or shortly before time-out ends but also the effect of reintroducing an explosive, out-of-control client back into the time-in setting. Reintroducing a misbehaving student to the time-in setting may bring about two undesirable outcomes: (1) the offending student may learn that regardless of his behavior, he will be able to rejoin his classmates after a short period, and (2) classmates predisposed to act out may realize that they, too, can escape but later rejoin the group without having to change their behavior. Time-out should not be terminated until such behaviors are under control, or an alternative form of behavior suppression can be executed.

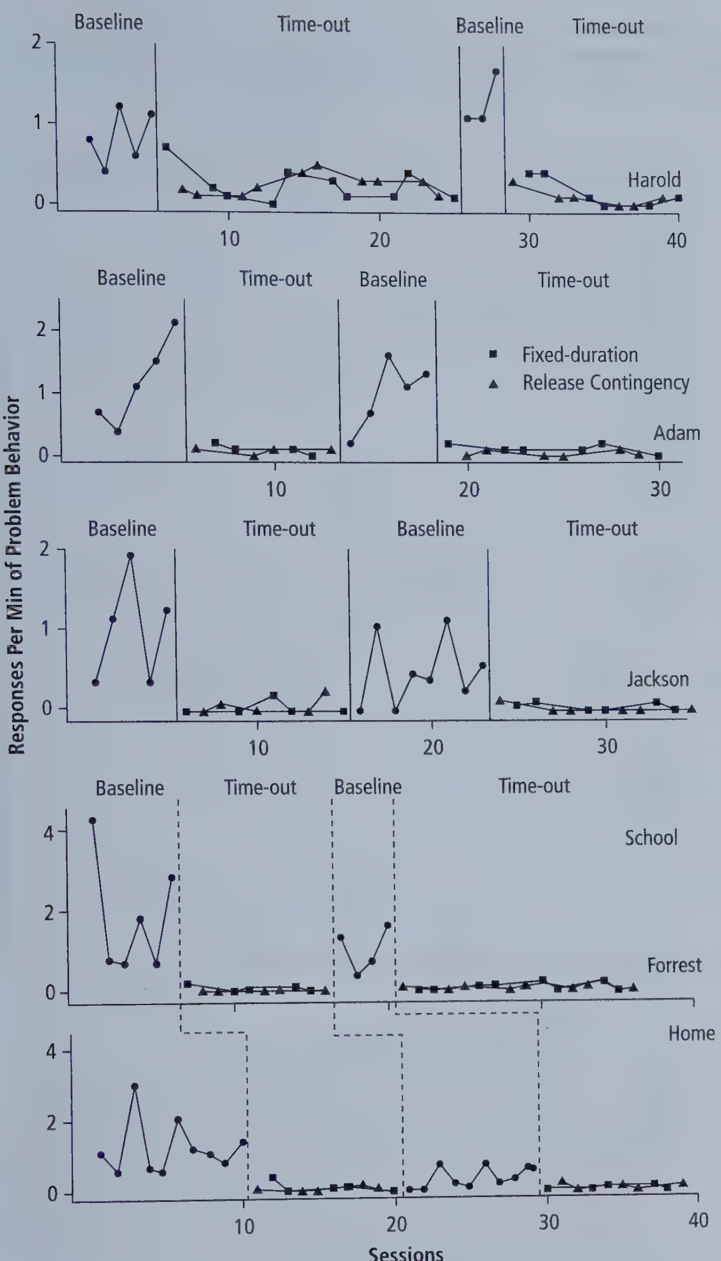


Figure 15.5 Rate of time-out–producing problem behavior across sessions for each participant.

"An Evaluation and Comparison of Time-Out Procedures with and without Release Contingencies", J. M. Donaldson and T. R. Vollmer, 2011, p, 699, *Journal of Applied Behavior Analysis, 44*. Reproduced with permission of John Wiley & Sons Inc.

If the practitioner anticipates that the inappropriate behavior that led to time-out may occur when time-out ends, then the individual should not be released from time-out. Exit criteria should establish this contingency, thereby increasing the potential that the problem behavior will not carry over to the time-in setting. Use of prompts, cues, or other discriminative stimuli toward the end of the time-out session may set the occasion for improved performance during the remaining minutes of time-out (e.g., "Rosario, when your time-out ends in 2 minutes, you may rejoin your classmates").

Ensure Compliance with Legal and Ethical Guidelines

The use of time-out in applied settings has been much discussed. Although litigation surrounding the use of this procedure has focused primarily on institutionalized populations (e.g., *Morales v. Turman*, 1973; *Wyatt v. Stickney*, 1974), the rulings in these cases have had a profound effect on the use of time-out in classroom, center, and clinical settings. The issues before the courts focus on the protection of client rights, whether time-out represents cruel and unusual punishment, and the degree of public acceptability of the time-out procedure.

The upshot of major court rulings has been that a person's right to treatment includes the right to be free from unnecessary and restrictive isolation. However, rulings also include language that permits the use of time-out in a behavior change program as long as the program is closely monitored and supervised and is designed to protect the individual or others from bodily harm.

In response to changing conditions in practice, court cases, increased knowledge about effective instruction, best practice, and input from practitioners, parents, and caretakers, the *Behavior Analyst Certification Board* (BACB) (2014) codified its view

Figure 15.6 Key distinctions between time-out, seclusion, and restraint.

	Definition	*Requirements*	*Compatible with Behavior Intervention Program (BIP)?*
Time-Out from Positive Reinforcement (Time-Out)	Immediate response-contingent withdrawal of the opportunity to earn positive reinforcers or the immediate loss of access to positive reinforcers for a specified time.	• A behavioral intervention plan (BIP) if exclusion time-out is used • Informed consent	• Yes, to the extent time-out is protective and therapeutic.
Seclusion	Involuntary isolation of the individual to a private, escape-proof setting, physically displacing the person from *time-in* to interrupt problem or harmful behavior.	• A BIP prohibiting use of a locked room & involuntary sedation • Informed consent • Alternative, evidence-based practices to be reviewed • Continuous monitoring	• Yes, to the extent seclusion is protective and therapeutic.
Restraint	Restricting the free movement of a person physically (e.g., basket hold) or mechanically (by a device) to prevent severe episodic, or self-injurious, behavior.	• Training on positive behavioral methods • De-escalation training • Regular safety, fire, and programmatic inspections • Continuous monitoring • Training to ensure health and welfare of the individual during restraint	• Yes, to the extent restraint is protective and therapeutic. • Yes, when used in an emergency situation to prevent immediate harm.

Based on The Council for Exceptional Children's (CEC, 2009) Policy on Physical Restraint and Seclusion Procedures in School Settings is an important source of guidance for school-based practitioners.

on ethical issues of punishment in its *Professional and Ethical Compliance Code for Behavior Analysts*. In this document, the BACB reaffirmed three guiding principles: (1) The welfare of the individual served is the highest priority, (2) individuals (and parents/guardians) have a right to choose, and (3) the principle of least restrictiveness should be followed. The BACB also provides its view on the use of punishment in applied settings in this document.

The Association for Behavior Analysis International's (ABAI) *Position Statement on Restraint and Seclusion* differentiates time-out, restraint, and seclusion in the context of serving clients with severe problem behavior. Figure 16.5 defines and summarizes key distinctions between these interventions.

Behavior analysts who serve clients with severe problem behavior should familiarize themselves with the entire ABAI position statement, which includes sections on client rights, clinical oversight, procedural safeguards, and ongoing evaluation based on objective data.

Applied behavior analysts who serve students with challenging behavior must also abide by policies and regulations governing the application of crisis interventions in schools. The Council for Exceptional Children's (CEC; (2009) *Policy on Physical Restraint and Seclusion Procedures in School Settings* is an important source of guidance for school-based practitioners (see https://www.cec.sped.org/~/media/Files/Policy/CEC%20Professional%20Policies%20and%20Positions/restraint%20and%20seclusion.pdf). Chapter 31 provides additional details on the ethical use of negative and positive punishment.

Evaluate Effectiveness

At the very least, data need to be obtained on the inappropriate behavior that initially led to time-out. If time-out is effective, the level of that behavior should be reduced substantially, and that reduction should be noticed by other persons in the environment. For legal and ethical reasons, additional records should be kept, documenting the use of time-out; the duration of each time-out episode; and the individual's behavior before, during, and after time-out (Yell, 1994).

Resources permitting, in addition to collecting data on the target behavior, it is sometimes beneficial to collect data on collateral behaviors, unexpected side effects, and the target behavior in other settings. For instance, time-out may produce emotional behavior (e.g., crying, aggressiveness, withdrawal) that might overshadow positive gains. Keeping a record of these episodes is helpful. Also, Reitman and Drabman (1999) suggested that for home-based applications of time-out, parents keep a record of the date, time, and duration of time-out, as well as a brief anecdotal description of the effects of time-out. When reviewed with skilled professionals, these anecdotal logs provide a session-by-session profile of performance, making adjustments to the time-out procedure more informed. In the view of Reitman and Drabman:

> The time-out record and discussions centered around it yielded a wealth of information including the type, frequency, context, and rate of misbehavior. . . . Graphing these

data provided visual feedback to both the therapists and the clients regarding the extent to which behavior change had occurred and facilitated the setting of new goals. (p. 143)

Terminate the Time-Out Program

Time-out has the potential to be employed well beyond its effective date. "Given the clinical and administrative effort required to initiate punishment programs, agencies may be inclined to keep them in place longer than needed rather than to expend more effort in evaluating their continued necessity" (Iwata, Rolider, & Dozier, 2009, p. 204).

Notwithstanding the general recommendation that the overall length of any behavior change program should be determined by data, the plan to eliminate time-out should begin shortly after the procedure is initiated. That is, once time-out has been introduced, and data show reduced levels of the target behavior, efforts should be made to switch to positive behavior reductive approaches (e.g., differential reinforcement of alternative behaviors, differential reinforcement of incompatible behaviors) or to reduce implementation over time. Conceptualizing time-out as existing on a continuum may give practitioners context for deciding for whom and under what conditions time-out is warranted. Stated in the form of questions, analysts might ask: Is time-out still necessary for this individual? Are there other less intrusive behavior reduction procedures that could be employed? If time-out is still considered the best method to reduce behavior, how can it be implemented less intrusively (nonexclusion versus exclusion)? Continuing time-out—or any punishment approach—beyond the point of its demonstrated effectiveness risks ethical violations and is counterproductive.

RESPONSE COST DEFINED

Response cost is the response-contingent loss of a specific number of positive reinforcers that has the effect of decreasing the future occurrence of similar responses. Essentially, response cost constitutes a fine for misbehavior.

Response cost typically involves the loss of one or more generalized conditioned reinforcers (e.g., money, tokens), tangibles (e.g., stickers, marbles), or minutes of already-earned time with a preferred activity (e.g., listening to music, playing a computer game) (Lee, Penrod, & Price, 2017; Musser, Bray, Kehle, & Jenson, 2001; Tanol, Johnson, McComas, & Cote, 2010; Wahl, Hawkins, Haydon, Marsicano, & Morrison, 2016). A teacher implements response cost by removing a specific number of positive reinforcers already held by a student contingent on the occurrence of a targeted inappropriate behavior.

The punitive value of the lost reinforcers depends on the individual reinforcement preferences of the participant. Hall and colleagues (1970), for example, found that the response-contingent removal of slips of paper bearing the student's name had a reductive effect on the number of complaints by a boy during 30-minute reading and math periods. Prior to each period during the response cost condition, the teacher placed five slips of paper bearing the boy's name on his desk. Each time the boy cried, whined, or complained during the period, the teacher removed a slip. A multiple baseline across academic periods with a reversal to baseline analysis revealed a functional relation between response cost and a marked decrease in disruptive behaviors (see Figure 15.7).

Desirable Aspects of Response Cost

Effectiveness

Response cost reduces the future occurrence of similar behaviors. According to Lerman and Toole (2011), ". . . the contingent removal of a specific amount of reinforcement (e.g., tokens) can function as an effective punisher" (p. 352). Boerke and Reitman (2011) underscore this point, citing the wide range of successful reductive response cost applications in home-, academic-, and hospital-based settings. Practitioners who face important challenges with disruptive,

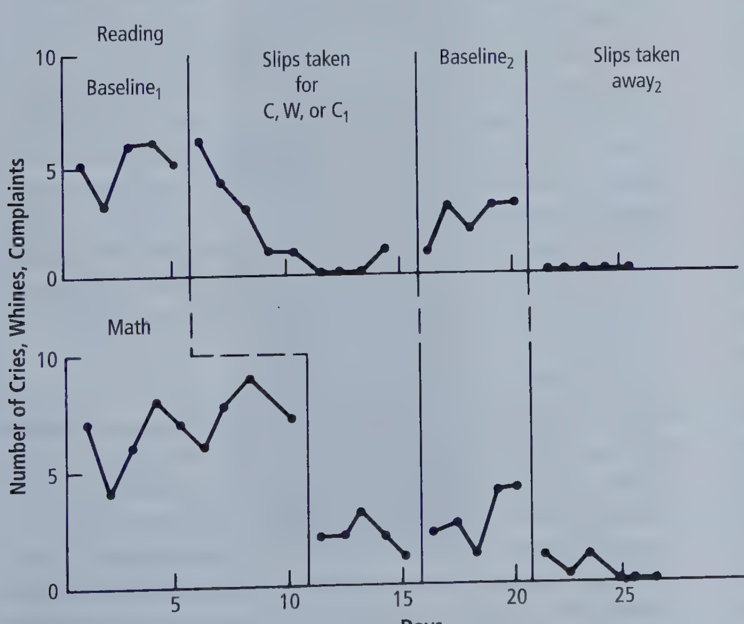


Figure 15.7 The number of cries, whines, and complaints by Billy during the 30-minute reading and arithmetic periods.

From "Modification of Disrupting and Talking-Out Behavior with the Teacher as Observer and Experimenter," by R. V. Hall, R. Fox, D. Williard, L. Goldsmith, M. Emerson, M. Owen, E. Porcia, and R. Davis, 1970, paper presented at the American Educational Research Association Convention, Minneapolis. Reprinted with permission.

off-task, non-compliant, and even life-threatening behaviors need a procedure that produces reliable reductive results. Response cost addresses this criterion.

Even though the data from numerous studies show the shorter term effectiveness of response cost, the longer term benefits for the individual, and the significant others with whom the individual interacts, are still largely unknown (Lerman & Toole, 2011). In their summary on developing function-based punishment procedures, of which response cost is a prime example, Lerman and Toole provide a number of recommendations for improving the long-term effectiveness of response cost:

> . . . punishment can be improved by (1) selecting punishers via pretreatment avoidance, choice, or activity assessments; (2) delivering the consequences immediately following each instance of the behavior; (3) ensuring that alternative sources of reinforcement are readily available; (4) establishing discriminative control . . . ; and (5) . . . using conditioned punishers. (pp. 363–364)

Practitioners would do well to follow Lerman and Toole's guidelines. Furthermore, applied behavior analysts charged with training field-based practitioners would also be wise to incorporate these five suggestions into their initial and recurrent response cost training programs. Chapter 31 contains additional strategies for trainers placed in supervisory positions.

Least Restrictive Alternative

We state repeatedly throughout this text, and especially in Chapter 31, that behavior analysts must comply with an ethical code of conduct. One important application of this code is that practitioners use interventions that are consistent with the principle of the least restrictive alternative as detailed in the Professional and Ethical Compliance Code for Behavior Analysts (2014). Briefly, this principle states that the most powerful, yet least intrusive, method be employed first and found to be ineffective, before more restrictive options are exercised. Hence, when analysts face challenging behaviors, especially those that have been resistant to other reductive treatment (e.g., extinction, differential reinforcement of lower rates of behavior), it is incumbent on the analyst to continue exploring least restrictive measures. Lerman and Toole (2011) add the following guidance.

> Ethical guidelines mandate that priority should be given to the least restrictive alternative procedure that is clinically effective. . . . Inherent in this approach to treatment selection is the need to arrange punishment procedures hierarchically based on the degree of restrictiveness (i.e., how much the procedure limits an individual's freedom or intrudes into an individual life in some manner) or aversiveness (i.e., how much the procedure produces discomfort, pain, or distress). Nonexclusionary time-out and response cost typically are considered the least restrictive of the procedures, followed by exclusionary time-out, overcorrection, and other physical punishers. (p. 353)

Moderate to Rapid Decrease in Behavior

Like other forms of punishment, response cost usually produces a moderate to rapid decrease in behavior. Unlike other behavior-reductive procedures that may require more time to show a significant effect (e.g., extinction, differential reinforcement of alternative behaviors), response cost tends to produce changes quickly. Suppressive effects are noted rapidly, within two to three sessions, as evidenced by Hall and colleagues' data (see Figure 15.7) and replicated by Donaldson, DeLeon, Fisher, and Kahng (2014) (see Figure 15.8).

For instance, when the response cost contingency was introduced during "Tokens" (denoted by the square data points in Figure 15.8), students lost one token per disruption, and the number of disruptive responses per minute fell to near zero. In a contrasting Token Earn condition, students could earn one token for emitting any appropriate behavior at the time a signal was emitted to record occurrences or nonoccurrences. This rapid loss was again noted after "Tokens: Choice" was reintroduced following Baseline 2. Overall, the data show an immediate and significant reduction in disruptive behavior. Interestingly, when students were given a choice (i.e., token earn or token loss), all four students shown in Figure 15.8, and two thirds of the other students in the entire study, preferred token loss over token gain.

Convenience

Response cost is convenient to implement and can be used in a variety of school-, home-, hospital-, and center-based settings (Ashbaugh & Peck, 1998; Musser et al., 2001). For example, if students are expected to follow explicitly stated classroom procedures, rules, or contracts, then any rule infraction will produce a loss of a specified reinforcer (e.g., a token), and it signals to the student that future occurrences of the misbehavior will result in the same consequence. Under these conditions, decision rules about implementation can be made expediently.

At home, response cost is convenient to the extent that it can be incorporated into an existing routine. For instance, if a child earns an allowance of $7 per week ($1 per day) for cleaning her room and placing her dishes in the sink after eating, noncompliance with one or both of these behaviors would lead to a monetary loss (e.g., 50¢ per infraction). In this setting, a response cost chart, posted in a visible location, would provide a convenient feedback mechanism to show the child and her parents accumulated losses (Bender & Mathes, 1995).

Figure 15.9 shows a sample response cost chart as it might be used in a classroom or tutorial setting. To implement the response cost procedure, the teacher writes a column of decreasing numbers on the board or chart. Immediately following the occurrence of a disruptive behavior, the teacher crosses off the highest remaining number. If the teacher places the numbers 15 through 0 on the board and three disruptive behaviors happened, the students would receive 12 minutes of preferred activity time that day.

Can Be Combined with Other Approaches

Response cost can be combined with other behavior change procedures (Athens, Vollmer, Sloman, St. Peter Pipkin, 2008; DeLeon et al., 2008; Long, Miltenberger, Ellingson, Ott, 1999; Watkins & Rapp, 2014). For instance, practitioners might add response cost to an existing treatment to further reduce levels of a target behavior. Researchers might add response cost to an existing procedure, withdraw it, and then reinstitute response

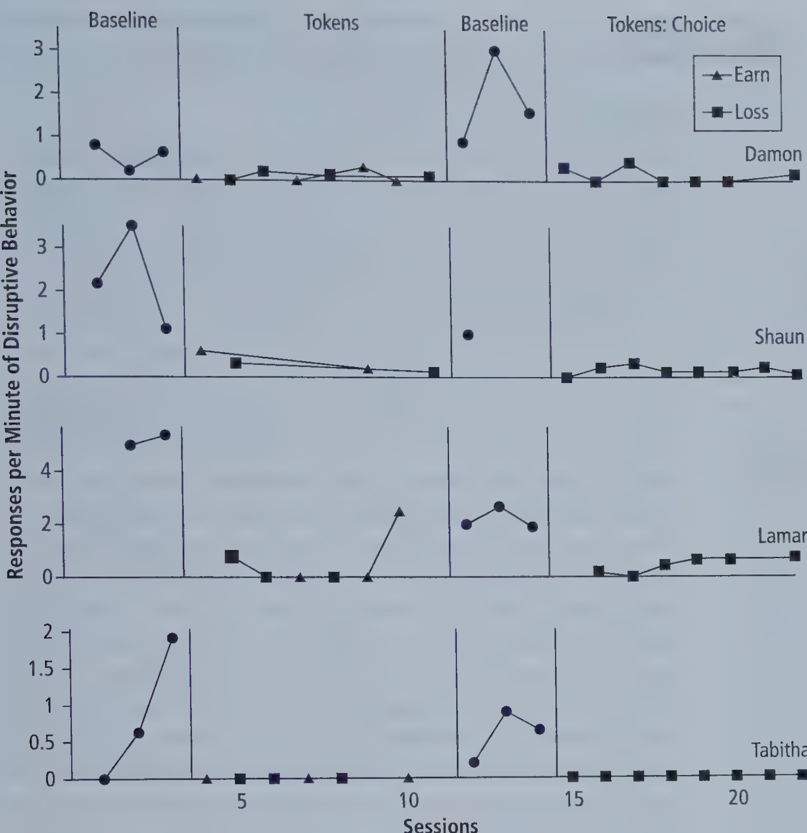


Figure 15.8 Responses per minute of disruptive behavior across sessions for Damon, Shaun, Lamar, and Tabitha.

"Effects of and preference for conditions of token earn versus token loss" by J. M. Donaldson, I. G. DeLeon, A. B. Fisher, and S. W. Kahng, 2014, *Journal of Applied Behavior Analysis*, p. 544. Reproduced with permission of John Wiley & Sons Inc.

cost, as part of a component analysis to calibrate the occurrence, rate, or percentage reduction. Such analysis might address important questions as to whether response cost alone is sufficient to produce desirable reductive levels, or whether it must be applied with another procedure to gain this effect. The following sections describe a few of these alternatives.

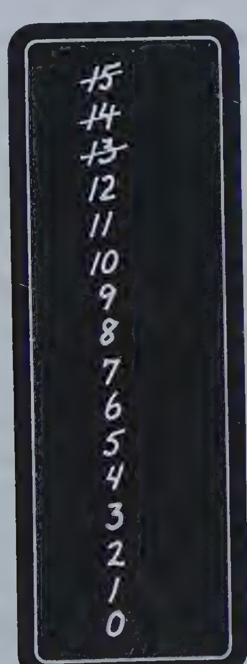


Figure 15.9 A response cost chart showing number of minutes of free time remaining after three rule infractions.

Response Cost with Positive Reinforcement. Response cost can be combined with positive reinforcement. For example, a student might receive a point, a star, or a sticker for each academic assignment completed during the morning seatwork period, but lose a point, a star, or a sticker for calling out without permission. Donaldson et al. (2014) used this dichotomous gain-loss procedure in their study.

Musser and colleagues (2001) combined response cost and positive reinforcement to improve the compliant behavior of three students with emotional disturbance and oppositional defiant behavior. Briefly, the procedures involved the teacher making a normal request for compliance (e.g., "Take out your book, please"). If the student complied within 5 seconds, the teacher delivered praise. If a 30-min period expired and no episodes of noncompliance were recorded, a sticker was delivered to the student. If the student failed to comply with a request within 5 seconds, the teacher waited 5 seconds and issued another directive (e.g., "Take out your book"). The teacher then waited an additional 5 seconds for compliance. If the student complied, the teacher issued praise. If not, one sticker was deducted for noncompliance. Stickers were later exchanged for a "mystery motivator" (e.g., a backup reinforcer from a grab bag). Results for all three students showed marked decreases in noncompliant behavior. Thus, a dual system, which was user friendly and convenient, proved to be effective insofar as the students earned stickers for intervals of compliant behavior, but lost stickers—the response cost contingency—for noncompliant behavior.

This method of combining response cost with positive reinforcement has at least two advantages: (1) If all of the tokens or points are not lost through response cost, the remaining ones can

be traded for backup conditioned or unconditioned reinforcers; and (2) the practitioner's behavior is not entirely focused on delivering a negative punishment contingency. Presumably, the practitioner will also focus on the occurrence of positive, constructive behaviors to reinforce. In doing so, the practitioner teaches new behaviors to the individual and reduces legal and ethical concerns.

Response Cost with Group Consequences. Response cost can be combined with group consequences. That is, contingent on the inappropriate behavior of any member of a group, the whole group loses a specific number of reinforcers. The Good Behavior Game, discussed in Chapter 28, serves as an excellent illustration of how response cost can be applied with groups.

Response Cost with Differential Reinforcement of Alternative Behavior (DRA). Response cost has been combined with differential reinforcement of alternative behavior (DRA) to treat the food refusal behavior of a 5-year-old child with developmental delay. Kahng, Tarbox, and Wilke (2001) replicated the Keeney, Fisher, Adelinis, and Wilder (2000) study, except that response cost was combined with DRA. During a baseline condition, various food group trials were presented every 30 seconds. If food was accepted, praise was delivered. If food was refused, the behavior was ignored. Based on reinforcer assessments, during DRA and response cost, books and audiotapes were made available at the meal. If food was accepted, praise was delivered. If food was refused, the books and audiotapes were removed for 30 seconds. The books and audiotapes were replaced during the next trial contingent on accepting food (i.e., DRA). Parents and grandparents were also trained in the DRA plus response cost procedures. Results showed a substantial increase in the percentage of intervals of food acceptance and a dramatic decrease in the percentage of intervals of food refusal and disruption.

Other Combinations. Watkins and Rapp (2014) combined response cost with environment enrichment to reduce the stereotypy behavior of five preadolescent and adolescent individuals with autism. In their study, participants were exposed to three conditions: no interaction; environment enrichment, during which reinforcers identified through a stimulus preference assessment were available; and environment enrichment, during which those preferred reinforcers were removed for 15 seconds contingent upon each occurrence of stereotypy behavior. Results showed that when environment enrichment and response cost were combined, the stereotypy behavior of all five individuals decreased immediately and markedly. Further, after these two procedures were withdrawn, the stereotypy behavior of three of the five participants did not increase.

In sum, the findings of studies directly combining response cost with other procedures have been positive. Whether response cost is used alone, or in combination with other procedures, it works. Still, practitioners should remember that (1) punishment procedures do not teach "new" behaviors; instead, they are designed to reduce the frequency of problem behavior; and (2) the environment where the target behavior occurs also needs to provide an ample number of reinforcers for appropriate behaviors. Stated differently, response cost should not be used as a sole behavior change approach without also considering the beneficial effects that accrue when positive reinforcement is provided. Combining response cost procedures with behavior-building approaches (e.g., DRA) is a good plan.

Established Social Validity

As discussed in Chapter 10, a procedure is considered socially valid if it meets three criteria: (1) The target behavior is significant for the individual, (2) the methodological procedures are acceptable and meet the standard of best care, and (3) the results produce meaningful positive change.

On the basis of these criteria, a legitimate claim can be made that response cost is socially valid. Typically, the behaviors selected for a response cost procedure call attention to themselves and have not improved in any meaningful way given prior interventions. Also, teachers and the students themselves report that response cost procedures are acceptable (Donaldson et al., 2014; Hirst, Dozier, & Payne, 2016). The dramatic reductive effect that response cost produces reaches the level of significance for those involved. That is, their lives are improved. Ultimately, they come into contact with more reinforcers and fewer punishers in the future.

Embry and Biglan (2008) included response cost in their list of 52 fundamental units of behavioral influence. In their view, response cost meets this designation because it is socially valid, is easy to implement, and produces clear and discrete outcomes that have been verified repeatedly by experimental and applied analyses.

RESPONSE COST METHODS

Response cost is implemented using two principal methods. The practitioner can (a) levy a fine against an individual's existing cache of previously earned reinforcers; or (b) levy a fine against an individual's surplus of extra, noncontingently provided reinforcers.

Existing Cache Response Cost

When a fine is levied against a person's existing cache of positive reinforcers, the individual loses a specific number of positive reinforcers from a supply of currently available reinforcers. For example, each occurrence of noncompliant behavior would produce a loss of 1 minute of free time from a current allocation of 15 minutes. In this case, response cost is applied to stimuli known to be reinforcing and accessible to the person.[3]

Bonus Response Cost

Bonus response cost operates when practitioners make additional reinforcers available noncontingently to the participant, specifically for removal when a target behavior occurs. Box 15.1 presents a vignette of how a father might implement a bonus response cost procedure with his two sons to reduce their fighting and squabbling at the dinner table.

USING RESPONSE COST EFFECTIVELY

To use response cost effectively, the behavior targeted for response cost, the timing of the fine, and the amount of the fine need to be explicitly stated. For instance, the behavior analyst

BOX 15.1

Father and Sons Vignette

Father: Boys, I'd like to talk to you.

Tom and Pete: OK.

Father: We have to find a way to stop the fighting and squabbling at the dinner table. The way you two go at each other upsets everyone in the family, and I can't eat my dinner in peace.

Tom: Well, Pete usually starts it.

Pete: I do not!

Tom: You do, too . . .

Father: (Interrupting) That's enough! This is just what I mean, nit-picking at each other. It's ridiculous and it needs to stop. I've given it some thought, and here's what we're going to do. Since each of you earns an allowance of $5 per week for doing household chores, I was tempted to take it away from you the next time either of you started squabbling. But I've decided against doing that. Instead, I'm going to make an additional $5 available to each of you, but for each sarcastic comment or squabble you get into at the dinner table, you'll lose $1 of that extra money. So, if you act properly, you'll get an additional $5 per week. If you have two squabbles, you'll lose $2 and earn only an additional $3. Do you understand?

Pete: How about if Tom starts something and I don't?

Father: Whoever does any fighting or squabbling loses $1. If Tom starts something and you ignore him, only he loses. Tom, the same thing holds true for you. If Pete starts something and you ignore him, only Pete loses. One more thing. Dinnertime starts when you are called to the table; it ends when you are excused from the table. Any other questions?

Tom: When do we begin?

Father: We'll begin tonight.

It is important to note that the father described the contingencies completely to his sons. He told them what would happen if no squabbles occurred—they would earn an additional $5 per week—and he told them what would happen if one sibling or the other started to bicker but the other ignored him. In addition, he clearly defined when dinnertime would begin and end. Each of these explanations was necessary for a complete description of the contingencies. Note that the money the boys normally earned as an allowance for doing household chores was not affected by this bonus response cost contingency. Ideally, the father praised his sons for their appropriate table manners and decorum when it occurred so that these behaviors could be reinforced and strengthened.

might say, "Any occurrence of an off-task behavior during reading will be subject to an immediate loss of one token." Further, any rules that would apply for refusal to comply with the response cost procedure need to be explained. However, in situations in which multiple behaviors are subject to the response cost contingency, or in which the degree of severity of the behavior determines the response cost, correspondingly greater fines should be associated with more severe behaviors. In effect, the magnitude of the punishment (response cost) should reflect the magnitude of the targeted behavior.

According to Weiner (1962)—the originator of the term *response cost*—the magnitude of the response cost fine is important. As the magnitude of the fine increases, larger reductions in the rate of the undesirable behavior may occur. However, if the loss of reinforcers is too great and too rapid, the person will soon be exhausted of his or her supply of reinforcers, and a deficit situation will exist. As a general rule, the fine should be large enough to suppress the future occurrence of the behavior, but not so large as to bankrupt the person or cause the system to lose its effectiveness. In short, if a client loses tokens or other reinforcers at too high a rate, he or she may give up and become passive or aggressive, and the procedure will become ineffective. Furthermore, fines should not be changed arbitrarily. If a teacher imposes a 1-unit token loss for each noncompliant response in the morning, the teacher should not impose a 3-unit token loss for the same behavior in the afternoon.

Deliver Fines Immediately

Ideally, the fine should be imposed immediately after the occurrence of each undesirable behavior. The more quickly response cost is applied subsequent to the occurrence of the behavior, the more effective the procedure becomes.

Response Cost or Bonus Response Cost?

Three considerations may help practitioners decide whether to initiate a response cost or a bonus response cost procedure. First, the least aversive procedure should be attempted initially. Consistent with the principle of the least restrictive alternative, an effort should be made to ensure that the minimum loss of reinforcers occurs for each infraction. Bonus response cost may be the less aversive of the two variations because the reinforcers are not deducted directly from the individual cache of reinforcers; instead, they are lost from a pool of noncontingently available (i.e., bonus) reinforcers.

The second consideration is similar to the first and can be stated in the form of a question: What is the potential for aggressive, emotional outbursts? From a social validity perspective, students (and their parents) would probably find it more agreeable to lose reinforcers from an available reserve than from their earnings. Consequently, a bonus response cost procedure may spark fewer aggressive or emotional outbursts, or be less offensive to students or their parents.

A third consideration is the need to reduce the behavior quickly. Combative or noncompliant behavior may be more

appropriately suppressed with response cost because the contingency directly reduces the student's available reinforcers. The response-contingent withdrawal of reinforcers serves to reduce the behavior swiftly and markedly.

Ensure Reinforcer Reserve

Positive reinforcers cannot be removed from a person who does not have any. Prior to using response cost, the practitioner must ensure a sufficient reinforcer reserve. Without such a reserve, the procedure is unlikely to be successful. For example, if a teacher used response-contingent withdrawal of free time for each occurrence of inappropriate behavior in a highly disruptive class, the students could exhaust all of their available free time before the end of the first hour, leaving the teacher wondering what to do for the remainder of the day. Deducting free time from time earmarked for succeeding days would hardly be beneficial.

Two suggestions apply to reduce the likelihood of having no reinforcers available. First, the ratio of points earned to points lost can be managed. If baseline data indicate that the inappropriate behavior occurs at a high rate, more reinforcers can be programmed for removal. Also, determining the magnitude of the fine, and stating it explicitly beforehand, is helpful. Minor infractions may warrant relatively smaller fines, whereas major infractions may warrant substantially higher fines. Second, if all reinforcers are lost and another inappropriate behavior occurs, the analyst might consider implementing time-out from positive reinforcement as a supplemental procedure.

To establish the initial number of available reinforcers, practitioners collect baseline data on the occurrence of inappropriate behavior during the day or session. The mean baseline figure can be increased for the number of reinforcers to ensure that all reinforcers are not lost when response cost is in effect. Although no empirically verifiable guidelines are available, a prudent approach is to increase the number of reinforcers 25% above the mean number of occurrences during baseline. For example, if baseline data indicate that the mean number of disruptions per day is 20, the practitioner might establish 25 minutes of free time (the positive reinforcer) as the initial level (i.e., 20×1.25). If the practitioner calculates points instead of percentages, she could add an additional 10 to 20 points to ensure an adequate buffer.

Recognize the Potential for Unplanned or Unexpected Outcomes

Two situations may require the implementation of a contingency plan. One occurs when the repeated imposition of response cost serves to reinforce, and not punish, the undesirable behavior. When this situation arises, the practitioner should stop using response cost and switch to another reductive procedure (e.g., time-out or DRA). The second situation occurs when the person refuses to hand over positive reinforcers. To reduce the likelihood of this event, the practitioner should clarify the consequences of such refusal beforehand and (a) make sure that an adequate supply of backup reinforcers is available, (b) impose an additional penalty for not giving up the reinforcer (e.g., the sticker) (Musser et al., 2001), (c) refund some fractional portion of the fine for complying with immediate payment, or (d) sidestep the whole possibility of refusal by maintaining control of the tokens on a ledger, electronically, or in a jar placed out of the participant's reach.

Avoid Overusing Response Cost

Response cost should be saved for those major undesirable behaviors that call attention to themselves and need to be suppressed quickly. A teacher's or parent's primary attention should always focus on positive behavior to reinforce; response cost should be implemented as a last resort and should be combined with other procedures that ultimately are designed to build adaptive behavior.

Keep Records

Each occurrence of response cost and the behavior that occasioned it should be recorded. Minimally, the analyst should record the number of times fines are imposed, the persons to whom fines are issued, and the effects of the fines. Daily data collection helps to determine the efficacy of the response cost procedure that might otherwise remain undetected (Boerke & Reitman, 2011). By graphing the effects of the response cost intervention, a chart is created that the behavior analyst can use to (1) determine the immediacy, the extent, and the duration of the suppressive effect of the procedure, and (2) document the effectiveness to a skeptical practitioner predisposed to using response cost based on biased personal beliefs, or inadequate preservice training (Cooke, 1984).

RESPONSE COST CONSIDERATIONS

Increased Aggression

Response-contingent withdrawal of positive reinforcers may increase student verbal and physical aggressiveness. The student who loses several tokens, especially within a short time, may verbally or physically assault the teacher. Emotional behaviors should be ignored whenever possible if they accompany the implementation of response cost (Walker, 1983). Still, teachers should anticipate this possibility and (a) preempt a decision to use response cost if they suspect that a worse condition will result in the aftermath of an emotional episode, or (b) be prepared to "ride out the storm."

Avoidance

The setting in which response cost occurs or the person who administers it can become a conditioned aversive stimulus. If this situation occurs in school, the student may avoid the school, the classroom, or the teacher by being absent or tardy. A teacher can reduce the likelihood of becoming a conditioned aversive stimulus by contingently delivering positive reinforcement for appropriate behavior.

Collateral Reductions of Desired Behavior

The response-contingent withdrawal of positive reinforcers for one behavior can affect the occurrence of other behaviors as

well (St. Peter, Byrd, Pence, & Foreman, 2016). If the teacher fines Angelina 1 minute of recess each time she calls out during math class, the response cost procedure may reduce not only her call-outs but also her math productivity. Angelina may say to the teacher, "Since I lost my recess time for calling out, I am not going to do my math." She could also become passive-aggressive by sitting in her seat with arms folded and not work. Teachers and other practitioners should anticipate such collateral behaviors and clearly explain the response cost rules, reinforce other classmates as models of appropriate behavior, and avoid face-to-face confrontations.

Calling Attention to the Punished Behavior

Response cost calls attention to the undesirable behavior. That is, upon the occurrence of an inappropriate behavior, the student is informed of reinforcer loss. The teacher's attention—even in the form of notification of reinforcer loss—could serve as an attention-getting reinforcing event. In effect, the teacher's attention may increase the occurrence of future misbehavior. For instance, the teacher may have difficulty because every mark that she places on the chalkboard indicates that a positive reinforcer has been lost (e.g., minutes of free time). To counteract the possibility of calling attention to inappropriate behavior, the teacher should ensure that the delivery ratio of reinforcers to the removal ratio of reinforcers favors positive reinforcement.

Unpredictability

As with other forms of punishment, the side effects of response cost can be unpredictable and are related to a number of variables that are not fully understood or well investigated across participants, settings, or behaviors. These variables include the magnitude of the fine, the previous punishment and reinforcement history of the individual, the number of times behaviors are fined, and the availability of alternative responses that are eligible for reinforcement.

Treatment Integrity

Treatment integrity is the accuracy and procedural consistency with which a planned intervention is executed. If the plan is not implemented with high fidelity, then it becomes extremely difficult to determine if the lack of demonstrated improvement is due to the intervention not being implemented as designed, or if the treatment itself was not robust enough to affect the change.

Response cost, like other behavior-building or -suppressing procedures, suffers from two possible threats to treatment integrity: omission errors and commission errors. Omission errors occur when the analyst fails to apply the response cost contingency and levy the appropriate fine after target behavior occurrence. In this situation, a target behavior occurs, and the practitioner either does not register the behavior as an occurrence or observes the behavior but still fails to apply the fine. In either case, an omission error occurs. Commission errors occur when practitioners levy fines for behavior that should not have been subjected to response cost. In other words, a conditioned generalized reinforcer is removed from the individual when it should not have been.

Three points are worth noting about omission and commission errors. First, providing supervisory training, practice, feedback, and recurrent maintenance instruction reduces the chances of these types of errors. Second, the higher the fidelity of treatment integrity—that is, approaching 100%—the more suppressed the behavior becomes. Finally, at least some data in nonclinical applications suggest that even if a response cost procedure is implemented partially incorrectly (e.g., a mix of omission and/or commission errors), acceptable suppression may still result (St. Peter et al., 2016). On the continuum of errors scale, obviously, practitioners should aim for zero, or few, errors, but be comforted by the fact that some errors can exist without the entire response cost program being severely compromised.

SUMMARY

Time-Out from Positive Reinforcement Defined

1. Time-out from positive reinforcement (or simply time-out) is the immediate response-contingent withdrawal of the opportunity to earn positive reinforcers or the immediate loss of access to positive reinforcers for a specified time.

2. To function as negative punishment, the time-out procedure must decrease future occurrences of similar behaviors. Implicit in the definition of time-out are three important factors: (a) The discrepancy between the "time-in" and the time-out environments must be discriminative, (b) the response-contingent loss of access to reinforcement must occur immediately, and (c) a resultant decrease in the future frequency of the time-out–producing behavior must occur.

Time-Out Tactics for Applied Settings

3. There are two basic types of time-out: nonexclusion and exclusion.

4. Nonexclusion time-out means that contingent on the occurrence of the target behavior, the participant remains physically within the time-in setting, but elements of that setting are changed so that (a) the opportunity to earn reinforcers is removed; (b) a specific reinforcer is terminated; (c) repositioning occurs so that observations of ongoing activity continue, but without reinforcement; or (d) the individual is removed to a different space within the time-in setting.

5. Nonexclusion time-out is typically implemented using one of four tactics: planned ignoring, terminate specific reinforcer contact, contingent observation, and partition/select space time-out.

6. Exclusion time-out is the physical separation of the participant from the time-in environment when time-out goes into effect.
7. There are two basic tactics for implementing exclusion time-out contingent on the occurrence of the target behavior: (a) The participant is removed immediately from the time-in environment, or (b) the time-in environment is removed immediately from the participant.
8. Time-out is a desirable alternative for reducing behavior because of its ease of application, acceptability, rapid suppression of behavior effects, and ability to be combined with other approaches.

Using Time-Out Effectively

9. The time-in environment must be reinforcing if time-out is to be effective.
10. Effective use of time-out requires that the behaviors leading to, the duration of, and exit criteria for time-out be explicitly stated. Further, practitioners must decide whether to use nonexclusion or exclusion time-out.
11. The exit criteria for time-out can be conceptualized in three dimensions: (a) the duration of any individual time-out session, (b) the behavioral criteria for leaving time-out and returning to the time-in setting, and (c) the overall length of the time-out program.
12. The duration of any individual time-out session, the behavioral criteria that prompt termination of time-out, and the duration of the time-out program need to be made explicit before launching time-out, and appropriate ethical and legal guidelines should be followed.
13. In the majority of applications, permission is required before time-out can be implemented.
14. Practitioners should be aware of legal and ethical considerations before implementing time-out. As a punishment procedure, it should be used only after positive reductive procedures have failed, and with planned monitoring, supervision, and evaluation considerations in place.

Response Cost Defined

15. Response cost is the response-contingent loss of a specific number of positive reinforcers that has the effect of decreasing the future occurrence of similar responses.
16. Response cost constitutes a fine for misbehavior.
17. Response cost typically involves the loss of one or more generalized conditioned reinforcers, tangibles, or minutes of already-earned time with a preferred activity.
18. The punitive value of the lost reinforcers depends entirely on the individual reinforcement preferences of the participant.
19. Response cost affords practitioners a number of desirable features. It is effective, meets the standard for least restrictive alternative, produces rapid suppression of behavior, is convenient, can be combined with other procedures, and is socially valid.

Response Cost Methods

20. Response cost is implemented using two principal methods. The practitioner can (a) levy a fine against an individual's existing cache of previously earned reinforcers; or (b) levy a fine against an individual's surplus of extra, noncontingently provided reinforcers.
21. When a fine is levied against a person's existing cache of positive reinforcers, the individual loses a specific number of positive reinforcers from a supply of currently available reinforcers.
22. Bonus response cost operates when practitioners make additional reinforcers available noncontingently to the participant, specifically for removal when a target behavior occurs.

Using Response Cost Effectively

23. To use response cost effectively, practitioners should deliver fines immediately, decide whether bonus response cost is a preferred option, ensure reinforcer reserve, recognize the potential for unplanned or unexpected outcomes, avoid overusing response cost, and keep records on its effects.

Response Cost Considerations

24. Implementing response cost may increase aggression, produce avoidance responses, affect collateral reductions of desired behaviors, call attention to the punished behavior, be unpredictable, and suffer treatment integrity challenges.

KEY TERMS

bonus response cost
contingent observation
exclusion time-out
negative punishment
nonexclusion time-out
partition or select space time-out
planned ignoring
response cost
terminate specific reinforcer contact
time-out from positive reinforcement (time-out)

MULTIPLE-CHOICE QUESTIONS

1. A punishment procedure has been put in place for Carlos. Whenever he swears in class his teacher says, "Carlos, swearing is not allowed in this classroom," and he is escorted to the time-out room where he has to stay for at least five minutes. His teacher has used the procedure for three weeks and his swearing has not decreased at all. What is the problem in this example?
   - **a.** The time-out room is not functioning as punishment.
   - **b.** Carlos should be placed in the time-out room for a longer period of time.
   - **c.** His teacher needs to explain to him why he is being placed in time-out.
   - **d.** His teacher has not done enough to reinforce the time-in setting.

   Hint: (See "Exclusion Time-Out")

2. Before implementing a contingent observation program for cutting in line at the drinking fountain, Mrs. Salkey made sure to send a note home to guardians telling them about the program and asking the guardians to contact her if there were any objections. In addition she checked with her principal to make sure she was following the schoolwide behavior management program. Which part of using time-out effectively does this example best exemplify?
   - **a.** Reinforce and enrich the time-in environment.
   - **b.** Define the behaviors leading to time-out.
   - **c.** Define the procedures for the duration of time-out.
   - **d.** Explain the time-out rules.
   - **e.** Evaluate effectiveness.
   - **f.** Obtain permission.

   Hint: (See "Using Time-Out Effectively")

3. When implementing a contingent observation program for cutting in line at the drinking fountain, Mrs. Salkey made sure to provide lots of tokens for students who waited their turn. Which part of using time-out effectively does this example best exemplify?
   - **a.** Reinforce and enrich the time-in environment.
   - **b.** Define the behaviors leading to time-out.
   - **c.** Define the procedures for the duration of time-out.
   - **d.** Explain the time-out rules.
   - **e.** Evaluate effectiveness.
   - **f.** Obtain permission.

   Hint: (See "Using Time-Out Effectively")

4. When implementing the contingent observation program, Mrs. Salkey kept a record of each student who cut in line every day for one week. By Friday, there were no students who cut in line. Which part of using time-out effectively does this example best exemplify?
   - **a.** Reinforce and enrich the time-in environment.
   - **b.** Define the behaviors leading to time-out.
   - **c.** Define the procedures for the duration of time-out.
   - **d.** Explain the time-out rules.
   - **e.** Evaluate effectiveness.
   - **f.** Obtain permission.

   Hint: (See "Using Time-Out Effectively")

5. Mrs. Salkey used tokens for waiting along with a contingent observation program. This approach best exemplifies:
   - **a.** Rapid suppression of behavior
   - **b.** Combined approaches
   - **c.** Ease of application
   - **d.** Acceptability

   Hint: (See "Desirable Aspects of Time-Out")

6. Mr. Marley docks the class one minute of extra recess time every time he has to say, "people quiet down please." This best exemplifies which response cost method?
   - **a.** Direct fines
   - **b.** Bonus response cost
   - **c.** Using a response cost with a group
   - **d.** Combining with positive reinforcement

   Hint: (See "Response Cost Methods")

7. In addition to using a response cost with an entire group, Mr. Marley is also using which procedure?
   - **a.** Direct fines
   - **b.** Bonus response cost
   - **c.** Using a response cost with a group
   - **d.** Combining with positive reinforcement

   Hint: (See "Response Cost Methods")

8. Mrs. Webb uses a response-cost procedure in which a student loses all points toward participation in a monthly pizza party any time he or she says "shut up" to another student. As a result, whenever students lose their points, they say "shut up" more often. What is the problem?
   - **a.** The magnitude of the fine is too great.
   - **b.** The points weren't functioning as reinforcers.
   - **c.** The students became more aggressive.
   - **d.** The fine should have been larger.

   Hint: (See "Using Response Cost Effectively")

ESSAY-TYPE QUESTIONS

1. As Type II punishment, response cost can be classified and is defined by function. Specifically, if the future frequency of the punished behavior is reduced by the response-contingent withdrawal of a positive reinforcer then a response cost has occurred. However, if the removal of the reinforcer increases the level of the behavior or has no effect on it, response cost has not occurred. This distinction between application and effect is the key aspect in the definition of response cost. Why is it important to make a distinction between application and effect?

   Hint: (See "Response Cost Defined")

2. Why is it important to reinforce functional alternatives to a problem behavior when using Type II punishment?

   Hint: (See "Using Time-Out Effectively" and "Response Cost Methods")

3. Suppose you are a third-grade teacher of a classroom with 25 students. Three of the students in the classroom are causing problems. The problems include failure to complete assignments and talking during independent seat-work. Pick a Type II punishment and pair it with a reinforcement program that you will use for the entire class. Please refer to Powell and Powell's (1982) time-out implementation checklist.

   Hint: (See "Using Time-out Effectively" and "Using Response Cost Effectively")

4. List the undesirable aspects of negative punishment and describe how one might minimize those aspects.

   Hint: (See "Response Cost Considerations")

NOTES

1. The term *time-out from positive reinforcement* first appeared in the behavioral literature in a monograph authored by Charles Ferster (1958). He described five experiments with chimpanzees and pigeons. "The basic paradigm used . . . 'time out' from positive reinforcement as the aversive event. To study the effect of the time out, a simple response, such as a pigeon's peck at a circular disk, or a chimpanzee's press of a toggle switch, was maintained when occasional pecks or presses were followed with food. To generate the aversive event, reinforcement was withheld in the presence of one stimulus but alternately continued in the presence of a second stimulus. When the rate of responding fell to zero in the presence of the stimulus where reinforcement was discontinued, that stimulus, called the time-out stimulus, could then be used as an aversive event and its effect on the base-line performance measured" (p. 1).
2. See the Association for Applied Behavior Analysis International's *Statement on Restraint and Seclusion, 2010.*
3. In some situations, removing unconditioned and conditioned reinforcers (e.g., food, free time) from a person is legally and ethically inappropriate. For instance, it would be inadvisable to remove minutes from mandated, structured leisure or free time using response cost. To avoid any potential problems, practitioners should consider another method, or they must obtain permission from the local human rights review committee by justifying the rationale for removing this reinforcer using response cost.

PART 6

Antecedent Variables

Parts 4 and 5 detailed the effects of various types of stimulus changes immediately following behavior. The two chapters in Part 6 address the effects of stimulus conditions and changes that take place prior to a behavior occurring. Behavior does not occur in an environmental vacuum. Every response is emitted in the context of a particular set of antecedent conditions, and these antecedent events play a critical role in motivation and learning.

What people do at any particular moment is partly a function of what they want at that moment. In Chapter 16, Jack Michael and Caio Miguel describe motivating operations, environmental variables that (a) alter the momentary effectiveness of some stimulus, object, or event as a reinforcer (or punisher); and (b) alter the current frequency of all behaviors that have been followed by that form of reinforcement.

Although the defining feature of reinforcement is an increase in the future frequency of behavior, reinforcement produces a second important effect. Chapter 17 details how stimuli immediately preceding a response, or present during reinforcement, acquire an evocative function over future occurrences of the behavior. The chapter describes how behavior analysts use differential reinforcement under changing antecedent stimulus conditions to achieve the desired degrees of stimulus discrimination and generalization.

CHAPTER 16

Motivating Operations

Jack Michael and Caio Miguel

LEARNING OBJECTIVES

- Define an establishing operation and provide characteristics of an establishing operation.
- Discuss the recent movement to the term motivating operation.
- Describe the difference between motivating operations and discriminative relations.
- Define unconditioned motivating operations.
- State various examples of UMOs as related to the human organism.
- Define conditioned motivating operations.
- State the different classifications of CMOs and how they operate.
- State various examples of CMOs as related to the human organism.
- Name some general implications for the use of motivating operations in the study of behavior analysis.

We (Jack and Caio) just love Mexican food, especially tacos. Of course, we do not think about Mexican food all the time, only around lunch- or dinnertime—in other words, when we have not had food for a while, or when we are hungry. However, sometimes, even if we have already eaten, just the sight of our favorite Mexican restaurant, or maybe even a TV ad about Mexican food, makes us want some. At those times, we look for a good Mexican restaurant and feel quite satisfied when we find one. When we get to the restaurant, the sight of tortilla chips makes us want salsa. We can eat salsa at any time, but it doesn't sound good unless we are eating some tortilla chips. After eating the chips and salsa, all we can think about is taking a sip of that ice-cold margarita. And when the chips are gone, finding the waiter to get some more, and to order our tacos, becomes the most important thing in the world.

The scenario above suggests that the value of certain stimuli may vary conditionally with other environmental factors. Mexican food is not wanted all the time, and behaviors that produce it (e.g., driving to the restaurant) vary depending on someone's motivation. From a behavior analytic perspective, if a certain behavior is not occurring, there is a chance that reinforcing consequences are not being produced. However, even if consequences are available, their value or effectiveness can vary over time. Thus, the concept of reinforcement alone cannot properly explain a person's motivation to act (Michael, 1993). Going to a restaurant and ordering food invariably produces food. However, the likelihood of going to a restaurant may vary depending on one's motivation to eat. When we say someone is motivated or wants to do something, we are assuming that (a) what is wanted would function as a reinforcer at that moment, and (b) the current frequency of any behavior that has previously been so reinforced will increase. This chapter discusses motivation from a behavior analytic perspective.

DEFINITION AND CHARACTERISTICS OF MOTIVATING OPERATIONS

Early in his career, Skinner (1938) used the term *drive* to describe changes in behavior that were a function of motivational variables.[1] Although the term appeared in Keller and Schoenfeld's (1950) didactic treatment of Skinner's work, these authors placed an emphasis on the specific environmental operations (i.e., deprivation, satiation, and aversive stimulation) that are necessary to establish the drive (Miguel, 2013). Although not conforming to the same usage set forth by Keller and Schoenfeld, the term **establishing operation (EO)** was reintroduced by Michael (1982, 1993) to describe any environmental variable that momentarily alters (a) the effectiveness of some stimulus, object, or event as a reinforcer; and (b) the current frequency of all behavior that has been reinforced by that stimulus, object, or event. Subsequently, Laraway, Snycerski, Michael, and Poling (2003) proposed the inclusive term **motivating operation (MO)** to encompass the fact that these variables serve not only to strengthen, but also to weaken, the value of a consequence as a reinforcer.[2] The term *setting event* has also been used to describe some of the variables that may fit under the MO definition (see Box 16.1: "What's a Setting Event?").

BOX 16.1

What's a Setting Event?

The term *setting event* appeared in the behavior analytic literature (e.g., Bijou & Baer, 1961; Kantor, 1959) to describe a set of variables that may affect stimulus-response relationships (e.g., McLaughlin & Carr, 2005). More specifically, this description pertains to variables that are not discrete, such as the presence/absence of immediate events or objects, deprivation/satiation, or recent social interactions, all of which may differentially affect specific functions of stimuli (Leigland, 1984; Wahler & Fox, 1981). An example would be sleep deprivation affecting rates of crying (Bijou & Baer, 1961).

The term *setting event* seems to describe the same set of environmental variables that we have classified as motivating operations (MOs). However, no clear definition or functional description of the term has appeared in the literature, and most descriptions of setting events are topographical rather than functional (Leigland, 1984; Nosik & Carr, 2015; Smith & Iwata, 1997).

The MO concept, on the other hand, describes any variable that has two specific effects, namely, value-altering and behavior-altering effects. Thus, "a greater evidential requirement exists to classify an event as an MO than to classify it as a setting event." (Nosik & Carr, 2015, p. 221). We agree with Nosik and Carr that the MO concept, with its functional definition and increasing evidence of its utility, is a better taxonomic option for describing the variables subsumed under the term *setting event*.

Thus, an MO can be defined as an environmental variable that has two effects: value-altering and behavior-altering effects. The **value-altering effect** is either (a) an increase in the reinforcing effectiveness of some stimulus, object, or event, in which case the MO is an establishing operation (EO); or (b) a decrease in reinforcing effectiveness, in which case the MO is an **abolishing operation (AO)**. The **behavior-altering effect** is either (a) an increase in the current frequency of behavior that has been reinforced by some stimulus, object, or event, called an **evocative effect**; or (b) a decrease in the current frequency of behavior that has been reinforced by some stimulus, object, or event, called an **abative effect**. In addition to frequency, other aspects of behavior such as response magnitude, latency, and relative frequency (occurrence per opportunity) can be altered by an MO. These relations are shown in Figure 16.1.

Food deprivation is an EO that *increases* the effectiveness of food as a reinforcer (value-altering effect) and *evokes* all behavior that has been reinforced by food (behavior-altering effect). In other words, after not having food for a while, food becomes a reinforcer, and behaviors such as looking for food are increased. Food ingestion (consuming food) is an AO that *decreases* the effectiveness of food as a reinforcer (value-altering effect) and *abates* all behavior that has been followed by food reinforcement (behavior-altering effect). In other words, after eating, the reinforcing effectiveness of food is decreased, and so is looking for food (along with other previously learned behaviors that have produced food). These changes of behavior are said to be momentary. For instance, food deprivation (MO) increases the value of food, as well as all behaviors that have led to food, for as long as the organism is deprived. As food is obtained, the effectiveness of food deprivation as an MO is decreased (the animal becomes less hungry), as evidenced by a reduction in (a) the value of food as a reinforcer and (b) food-seeking behaviors.

Although the term *hunger* has been used to describe the effects of food deprivation, the term is simply a description (i.e., tact) of physiological changes that may accompany food deprivation (e.g., hunger pangs). Although these physiological sensations may prompt someone to get food, deprivation operations have direct value-altering and behavior-altering effects (Miguel, 2013). Thus, terms such as *hunger* and *thirst* are unnecessary (intervening) variables that may obscure a behavior analysis of motivation. The effects of the MO related to food are shown in Figure 16.2.

Painful stimulation is an EO that *increases* the effectiveness of pain reduction as a reinforcer and *evokes* all behavior that has been reinforced by pain reduction. So, a headache would make its reduction (or elimination) an effective reinforcer, and would also evoke behaviors, such as taking an aspirin, that have produced headache removal in the past. Of course, these two effects would only be observed while the organism is experiencing a headache, and for this reason we refer to the effects of the MO as being momentary. Conversely, a decrease in painful stimulation is an AO that *decreases* the effectiveness of pain reduction as a reinforcer and *abates* all behavior that has been followed by pain reduction. In other words, once the headache is gone, its elimination no longer functions as a reinforcer, and the likelihood of taking an aspirin is greatly reduced. The effects of the MO related to painful stimulation are shown in Figure 16.3.

Most of the statements in this chapter about value-altering and behavior-altering effects refer to relations involving reinforcement rather than punishment. It is reasonable to assume that MOs also alter the effectiveness of stimuli, objects, and events as *punishers*, with either establishing or abolishing effects; and also alter the current frequency of all behavior that has been *punished* by those stimuli, objects, or events, either abating or evoking such behavior. Later in the chapter we will consider MOs for punishment.

Isolating Value- and Behavior-altering Effects

The value-altering effect of an MO influences the potency of a reinforcer to increase the future frequency of behavior (Laraway, Snycerski, Olson, Becker, & Poling, 2014). An MO's

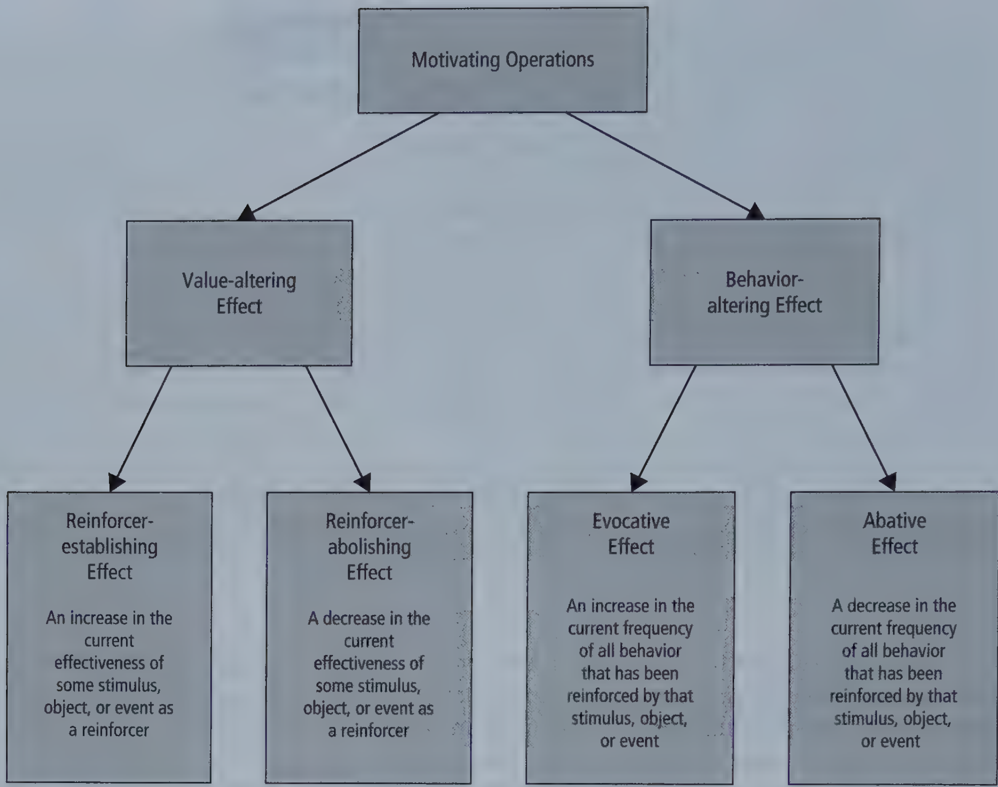


Figure 16.1 Motivating operations and their two defining effects.

value-altering effect is observed when the manipulation of certain environmental variables (e.g., 4, 8, and 12 hours of food deprivation) affects the speed with which an organism acquires new behaviors that produce the relevant consequence (e.g., food). In this way, an MO directly changes the potency of reinforcers (or punishers) to increase and maintain (or decrease and eliminate) behavior (Laraway et al., 2014).

The value of a reinforcer may also vary depending upon specific response requirements (DeLeon, Frank, Gregory, & Allman, 2009). One way to assess the potency of a specific reinforcer is to use a progressive ratio (PR) schedule in which response requirements increase after the delivery of each reinforcer (Roane, 2008; see Chapter 13). Under this arrangement, an organism may respond under high schedule requirements to obtain one consequence, but not another, even when both consequences have been observed to increase behavior under low response requirements (Roane, Lerman, & Vorndran, 2001). This procedure may serve to measure the momentary value of a reinforcer whose effectiveness varies depending on the MO. Other potential measures of reinforcer potency may include choice preparations, and latency to respond during a discrete trial (Poling, Lotfizadeh, & Edwards, 2017).

The behavior-altering effect is interpreted as a change in frequency due to the organism's behavior contacting reinforcement, implying that this change occurs only *after* the reinforcer has been obtained. However, a strong relation exists between the level of the MO and responding *under extinction*—when no reinforcers are being received (Keller & Schoenfeld, 1950). Early research (Skinner, 1938) showed that rats that had learned to press a lever to obtain food, when exposed to an extinction condition at different levels of deprivation, pressed the lever at different rates according to their levels of deprivation. Thus, different deprivation levels differentially evoked responses that had been previously learned. Similarly, if a child's tantrum has been followed by a parent's attention, the duration and intensity of the tantrum may be correlated with the amount of attention that the child has received before the tantrum.

The value-altering and behavior-altering effects of an MO occur simultaneously but are considered independent in the sense that one does not derive from the other. Food deprivation, for instance, increases behaviors that have been previously reinforced by food *prior to obtaining any food* (evocative effect). Food deprivation would also increase the reinforcing potency of food (the establishing effect). Thus, food-seeking behaviors would increase further after contacting food (Michael, 2000). Unfortunately, once behavior contacts reinforcement, these two effects may become difficult to discern (Poling et al., 2017).

Direct and Indirect Effects of Motivating Operations

The behavior-altering effect of an MO can increase (or decrease) the frequency of behavior because of a direct evocative (or abative) effect. The strength of the MO seems to be correlated with its capacity to directly evoke previously learned behavior. For example, when food deprivation is high, food-seeking behaviors may occur in situations that differ from the situations in which those behaviors were reinforced in the past (Lotfizadeh, Edwards, Redner, & Poling, 2012). An individual who has not eaten for a few hours is likely to request a snack from an acquaintance who usually carries some in her backpack (S^D). However, under extreme food deprivation, the individual might

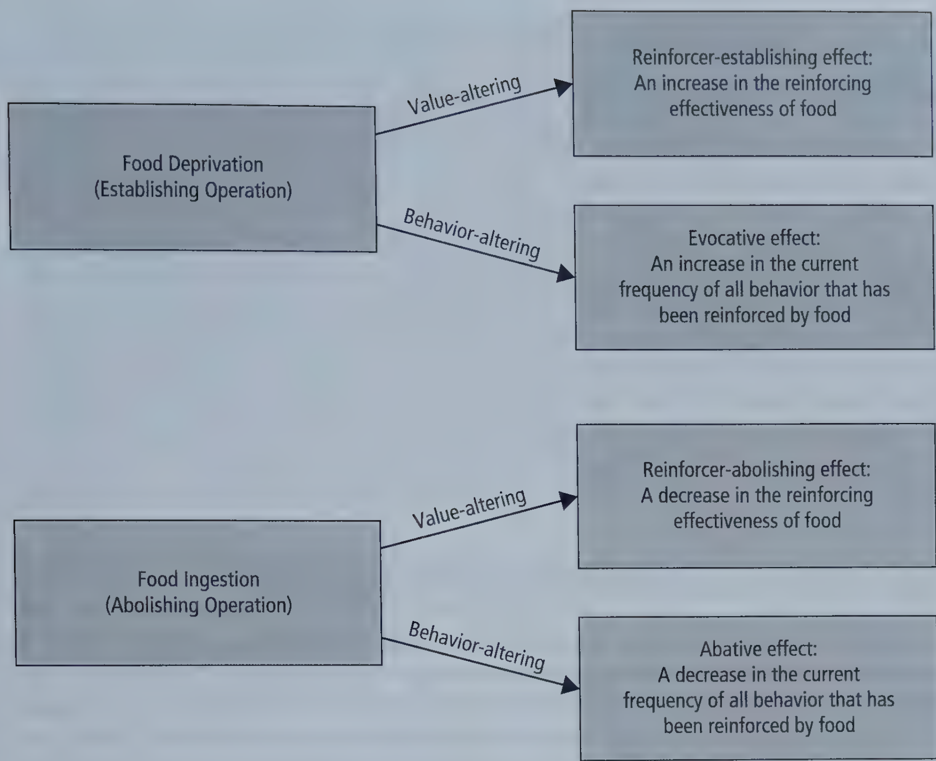


Figure 16.2 Motivating operations related to food.

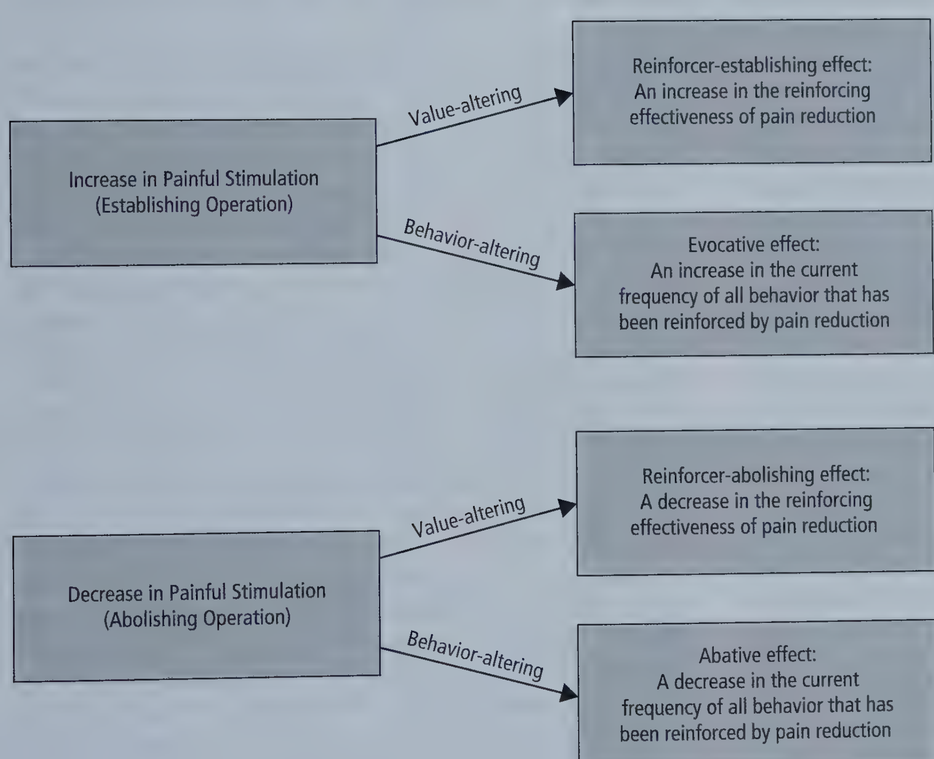


Figure 16.3 Motivating operations related to painful stimulation.

ask strangers for a snack, even though there has been no previous history of obtaining food from them. It seems that "the range of stimuli that evokes responding is wider when deprivation is high than when it is lower" (Lotfizadeh et al., 2012, p. 98). In other words, strong MOs can directly evoke behavior in the absence of relevant discriminative stimuli or, at least, in the presence of untrained stimuli. Thus, the direct evocative effect of the MO may be described as its capacity to evoke behavior irrespective of relevant discriminative stimuli.[3] It is important to remember that individuals do not need to be "aware" of the effects of reinforcement for it to affect behavior. Likewise, the effectiveness of MOs does not depend on individuals being able to describe the effects of MOs upon their behavior (see Box 16.2, "Do You Need to Know You Are Motivated?").

BOX 16.2

Do You Need to Know You Are Motivated?

The variables responsible for behavior-altering effects are often misconstrued by assuming that individuals must understand (verbally describe) the situation, and then behave appropriately as a result of this understanding. It is safe to assume that when pressing a lever to obtain water, a rat does not understand that lever presses have been reinforced by water in the past, nor that it should press the lever when water deprived. Similarly, humans do not need to verbally describe the situation for the MO to affect behavior. The fact that the effects of reinforcement upon behavior are automatic suggests that the behavior that has led to reinforcement will be momentarily increased when the relevant MO is in effect. Thus, just as the effects of reinforcing consequences do not depend on the organism's understanding of the contingency, neither do the effects of the MO upon behavior require that the organism be able to describe the relationship between behavior and consequence. Someone who spends a lot of time studying may not be able to identify the contingencies that have led to the establishment of this behavior or serve to momentarily evoke it. The frequency (or duration) of studying is still a function of environmental variables, whether the individual can describe them or not.

The MO indirectly affects the evocative or abative strength of relevant discriminative stimuli (S^Ds). Food deprivation, for instance, would increase the discriminative control of any stimulus that has been correlated with the availability of food. Therefore, the friend who carries snacks in her backpack would serve as a stronger S^D that would evoke a request when food deprivation is high. Also, the stronger the headache (MO), the more likely it is that the sight of the aspirin (as an S^D) would evoke the behavior of taking it. Note that the value of reinforcers present during discrimination training also depends upon a current MO (Laraway et al., 2003). In other words, the establishment of stimulus control over behavior not only depends on but also varies as a function of MOs. For instance, a child would never learn to discriminate between blue and red if no *effective reinforcer* were available for correct responses. To summarize, in the absence of an MO that establishes the value of the consequence as a reinforcer, discrimination learning is unlikely to occur, and a previously effective S^D would also not evoke behavior.

The value-altering effect also indirectly influences the reinforcing effectiveness of relevant conditioned stimuli. For example, food deprivation not only increases the capacity of food as a consequence to strengthen behavior (direct establishing effect), but also enhances the reinforcing effectiveness of all other stimuli that have been correlated with food, such as the smell of food, the sight of a restaurant, and so on. In the same way, a headache would not only establish its removal as an effective reinforcer, but also all stimuli that have been paired with pain reduction, such as an aspirin, a nap, a dark environment, and so forth. The latter effect, a conditioned one, will be discussed in the section on conditioned MOs, below. Figure 16.4 displays the different direct and indirect effects of the MO.

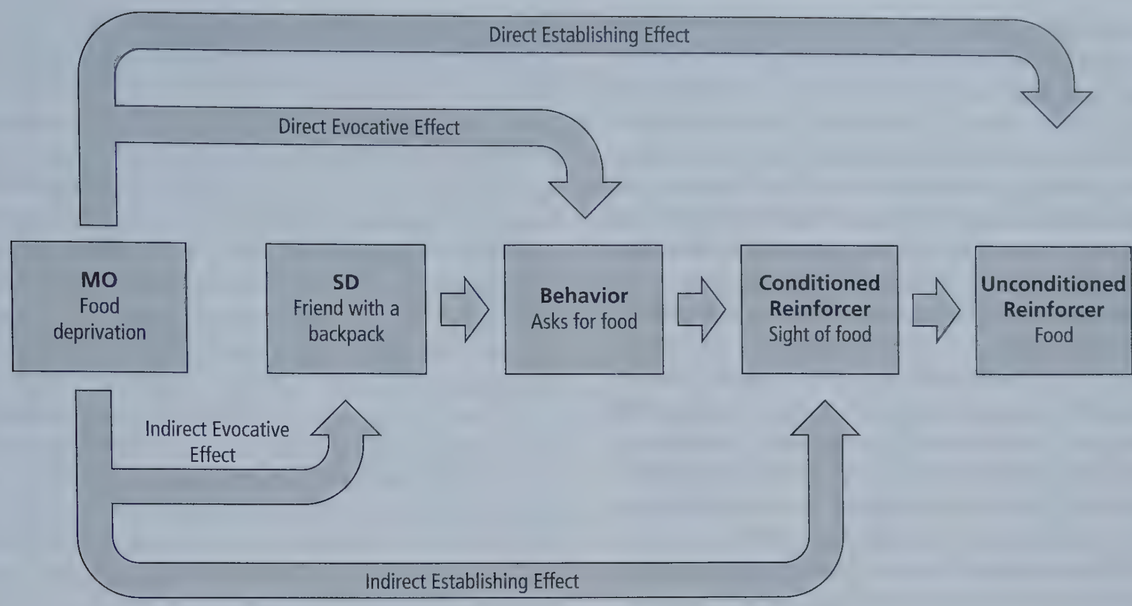
Behavior-altering versus Function-altering Effects[4]

Both MOs and S^Ds are antecedent variables that have behavior-altering effects. Antecedents can evoke or abate responses, but their occurrence does not alter the organism's functional behavioral repertoire. Thus, as antecedent variables, MOs alter the *current* frequency of all behaviors relevant to that MO. For example, a difficult task may serve as an establishing operation that momentarily increases all behaviors that in the past have produced removal of the task. Presenting a difficult task to a child who has learned that tantrums often result in task postponement or removal is likely to produce a tantrum (i.e., evocative effect). Conversely, if the task is not presented, or made easier (an abolishing operation), then the likelihood of tantrums is diminished (i.e., abative effect). However, as an intervention strategy, the removal of the task does not alter the child's repertoire; it only momentarily abates behavior that would otherwise occur if the motivating variable (difficult task) were in place. In other words, the child will not behave differently in the presence of difficult tasks solely as a result of antecedent manipulations.[5] Thus, reducing or eliminating the MO for a specific problem behavior is a temporary rather than permanent solution. When the MO is again in place, the problem behavior may re-occur (Michael, 2000). Nonetheless, eliminating environmental conditions (e.g., MOs) that contribute to the occurrence of problem behavior may be necessary,[6] especially when these conditions are impoverished or aversive (McGill, 1999).

Consequences can change the organism's repertoire, enabling the organism to behave differently in the future. Thus, consequences have **function-altering effects**. These variables include reinforcers, punishers, and the occurrence of a response without its reinforcer (extinction procedure), or without its punisher (recovery from punishment procedure). Consequences alter the *future* frequency of whatever behavior immediately preceded those consequences. In the previous example, if task removal or reduction of task difficulty is made contingent upon the child asking for a break or help, the frequency of these behaviors would increase under similar conditions in the future. In other words, the child's functional repertoire is changed as he or she can engage in two new responses: asking for a break or asking for help.

DISTINGUISHING BETWEEN MOs AND S^Ds

As described above, MOs and S^Ds are both antecedent variables that alter the current frequency of some particular behavior (i.e., evoke). Each operant variable influences response frequency because of its relation to reinforcing or punishing



Food deprivation directly establishes food as a reinforcer (value-altering effect) and evokes food-seeking behaviors (behavior-altering effect). Food deprivation indirectly establishes the effectiveness of all conditioned reinforcers correlated with food (value-altering effect) and indirectly affects the salience of the discriminative stimulus correlated with food (behavior-altering effect).

Figure 16.4 Direct and indirect effects of a motivating operation (MO).

consequences. An S^D affects behavior because its presence has been correlated with the *differential availability* of an effective reinforcer in the past. Differential availability means the relevant consequence has been available in the presence of the S^D (i.e., reinforcement) and unavailable in its absence (i.e., extinction; see Chapter 24).

For instance, if a food-deprived rat receives food for lever presses in the presence, but not absence, of a light, the light will eventually evoke lever presses. In this case, the light would be functioning as an S^D, since food delivery was more available in its presence than in its absence. Note that although food was never delivered when the light was off (the S^Δ condition), food would have served as an effective reinforcer had it been delivered when the light was off. In other words, food was valued (i.e., reinforcing) in both conditions, but available only when the light was on. The reduced frequency of lever presses during the S^Δ condition is due to extinction.

An MO, in contrast, controls behavior because of its relation to the *differential effectiveness* of a reinforcer for that behavior. Differential effectiveness means that the relevant consequence has been effective in the presence of, and ineffective in the absence of, the MO. In the example above, the rat would press the lever only in the presence of the light (S^D) when the MO is present (i.e., food deprived) and not when the MO is absent (i.e., food satiated). When the MO is absent, behavior decreases not because the reinforcer is unavailable (extinction), but because it is no longer valuable or effective.

Now imagine that the rat learns to press the lever to eliminate shock. In this case, pressing the lever produces pain reduction (or elimination) as a form of reinforcement. If painful stimulation as an antecedent were to be considered an S^D, pain reduction (the reinforcer) must have been available as a consequence in the presence of painful stimulation, but not in its absence. More importantly, like food in the previous example, pain reduction would have been effective as a reinforcer in both conditions, regardless of its availability. It is with respect to this requirement that some motivating variables fail to qualify as discriminative stimuli. Pain reduction could not serve as a reinforcer in the absence of pain as an antecedent.[7] In other words, in the absence of painful stimulation, behavior has not decreased due to extinction, but because there is no possible valuable reinforcer. The organism is not motivated to engage in behavior to reduce pain, if no pain is being felt.

In the previous example of light-on/light-off discrimination training, the reinforcer is available in the light-on condition and unavailable, yet still effective, in the light-off condition. In other words, the rat still "wants food" when the light is off but food is not available. The behavior occurs less frequently in the light-off (S^Δ) condition because it no longer produces an effective reinforcer. In contrast, in the absence of food deprivation or painful stimulation, food or pain removal, respectively, is not an effective reinforcer. Hence, food deprivation and painful stimulation qualify as MOs because they momentarily alter (a) the effectiveness of some stimuli, objects, or events as reinforcers (value-altering effect); and (b) the frequency of the behavior that has been reinforced by those stimuli, objects, or events (behavior-altering effect).

Similarly, an itch does not serve as a discriminative stimulus that evokes scratching it. Although the reinforcer (removing the itch) is available in the presence of the itch and unavailable in the absence of the itch, this is not an effective reinforcer in the absence of the itch. In this sense, the itch is best interpreted as functioning as an MO that establishes its removal as an effective reinforcer, and momentarily evokes scratching. In other words, having an itch is what motivates you to scratch.

Many antecedent stimuli that have been considered to evoke behavior as S^Ds would be better classified as MOs. This is particularly the case with aversive stimuli. For example, a headache is an MO that directly establishes headache removal as an unconditioned reinforcer and indirectly establishes aspirin as a conditioned reinforcer (see Figure 16.5). A headache also evokes behaviors that have produced pain reduction (finding and taking aspirin) in the past. If aspirin is found in the bathroom but not in the kitchen, then the bathroom will function as an S^D for the behavior of looking for aspirin because it has been correlated with *availability* of an *effective* reinforcer, while the kitchen will function as an S^Δ for looking for aspirin because it has been correlated with the *unavailability* of an *effective* reinforcer. In other words, for discrimination to take place, an individual needs to find aspirin in the bathroom, and not find aspirin in the kitchen, respectively, when needed (MO; see first two rows of Figure 16.5). In the absence of the headache (AO), one can still find aspirin in the bathroom, but there is no reason to look for it (available, but not a reinforcer), since there is no headache to be removed (unavailable non-reinforcer). Thus, in the absence of the headache, looking for an aspirin will not occur in either location (see bottom row of Figure 16.5).

In applied settings, when a child is given an instruction to complete a difficult task, the reinforcer may be task completion or termination. Additionally, the instruction would also acquire conditioned aversive properties due to its correlation with difficult tasks. Thus, the instruction could serve as an MO that establishes its removal as a reinforcer and evokes avoidance behavior (Carbone, Morgenstern, Zecchin-Tirri, & Kolberg, 2010). In contrast, if task completion has produced praise as a reinforcer in the presence, but not the absence, of the instruction, then the instruction may be functioning as an S^D. Other events best classified as MOs rather than S^Ds will be discussed later in this chapter, most specifically in the section on conditioned motivating operations.

Although behavior analysts use the three-term contingency (antecedent-behavior-consequence) to understand operant behavior, any given response may be evoked by a multitude of variables functioning as antecedents (i.e., multiple control). A child who enters the kitchen and asks for a cookie does so because in the past such behavior has produced reinforcement under similar circumstances. These circumstances involve variables that establish cookies as reinforcers, as well as increase the likelihood of asking for them (food deprivation as an MO). However, the behavior is more likely to occur in environments where cookies are available (kitchen as an S^D) and in the presence of an adult (also an S^D) who can give the child a cookie (Skinner, 1957). In this case, food deprivation is correlated with increasing the effectiveness of cookies as reinforcers, while the kitchen and the adult are both correlated with the availability of cookies.

To summarize, S^Ds alter the differential availability of a currently effective form of reinforcement for a particular type of behavior; MOs alter the differential reinforcing effectiveness of a particular type of environmental event. In nontechnical terms, an S^D tells you that something you want is available; an MO makes you want something.

UNCONDITIONED MOTIVATING OPERATIONS (UMOs)

For all organisms there are events, operations, and stimulus conditions with unconditioned *value-altering* effects. In other words, these MOs establish or abolish the value of stimuli

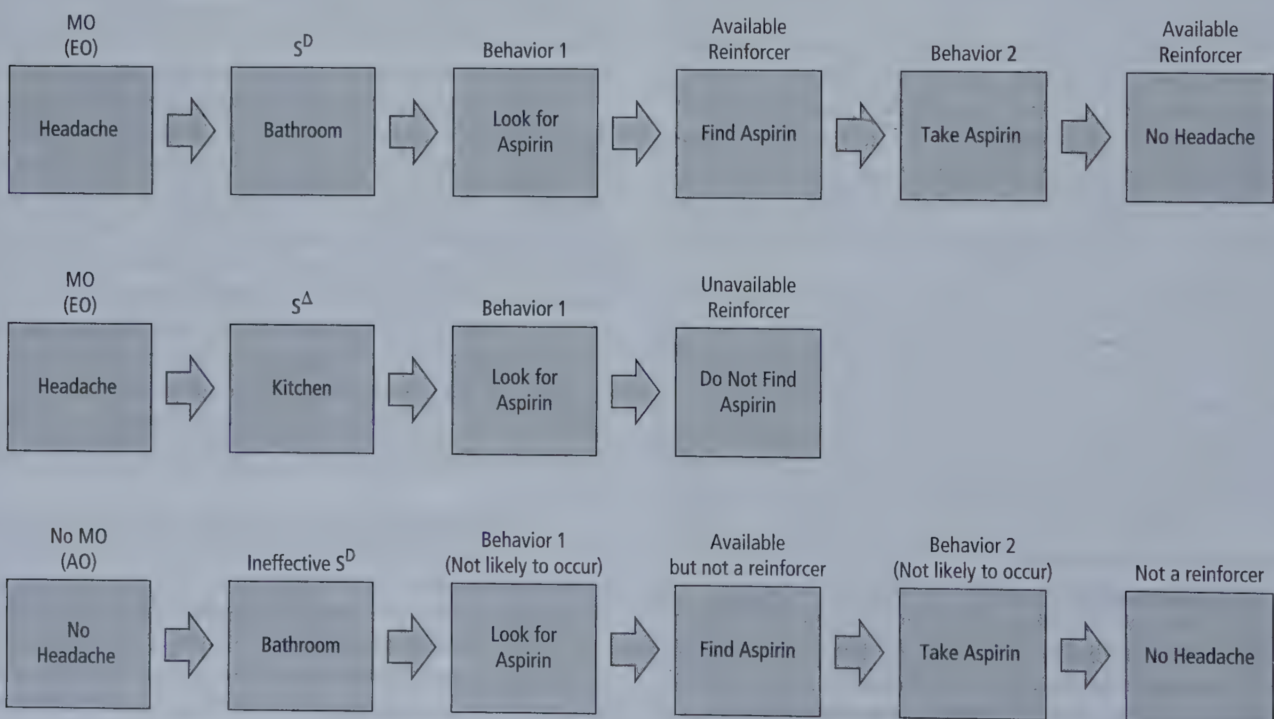


Figure 16.5 Distinguishing between S^D and MO.

as reinforcers in the absence of prior learning. For example, humans are born with the capacity to be affected by food reinforcement as a result of food deprivation and by pain reduction reinforcement as a result of pain onset or increase. Thus, food deprivation and painful stimulation are **unconditioned motivating operations (UMOs)**.[8]

MOs are classified as unconditioned based upon the unlearned aspect of their value-altering effects, since the behavior-altering effects of MOs are usually learned. Said another way, we are born with the capacity to be affected by food reinforcement as a result of food deprivation, but we learn the behaviors evoked by food deprivation—asking for food, going to where food is kept, cooking, and so forth.

Deprivation of food, water, oxygen, activity, and sleep all have **reinforcer-establishing** and **evocative effects** (see Table 16.1). For instance, water deprivation momentarily establishes the effectiveness of water as a reinforcer and evokes all behaviors that have produced water. By contrast, food and water ingestion, oxygen intake, being active, and sleeping have **reinforcer-abolishing** and **abative effects** (see Table 16.2). For instance, water ingestion abolishes the effectiveness of water as a reinforcer and abates all behaviors that have produced water.

The passage of time since last sexual activity functions as a UMO relevant to sexual stimulation. This form of deprivation establishes the effectiveness of sexual stimulation as a reinforcer and evokes behavior that has produced such stimulation. In contrast, sexual orgasm is a UMO that abolishes (weakens) the effectiveness of sexual stimulation as a reinforcer and abates (decreases the frequency of) behavior that has achieved that kind of reinforcement. In addition, tactile stimulation of erogenous regions of the body seems to function as a UMO in making further similar stimulation even more effective as a reinforcer, and in evoking all behavior that in the past has achieved such stimulation.

UMOs related to temperature changes include becoming uncomfortably cold and uncomfortably warm. Getting cold is a UMO that establishes getting warmer as a reinforcer and evokes behavior that has produced warmth, such as putting on a jacket or moving closer to a heat source. A return to a comfortable (i.e., nonaversive) temperature is a UMO that abolishes increased warmth as a reinforcer and abates behavior that has produced warmth. In this case, behaviors such as putting on a jacket are less likely to occur. Getting uncomfortably warm is a UMO that establishes a cooler temperature as an effective reinforcer and evokes behaviors that have resulted in a body-cooling effect. For example, getting warm may evoke the behavior of turning on the air conditioning. A return to a normal temperature abolishes being cooler as a reinforcer and abates body-cooling behavior. In this case, turning on the air conditioning is less likely to occur.

An increase in painful stimulation establishes pain reduction as a reinforcer and evokes the (escape) behavior that has achieved such reduction. As described previously, a headache establishes its own removal as a reinforcer and evokes any behavior that has eliminated headaches, such as taking aspirin. A decrease in painful stimulation abolishes the effectiveness of pain reduction as a reinforcer and abates the behavior that has been reinforced by pain reduction. Thus, when a headache is gone, headache removal is no longer a reinforcer and taking an aspirin is less likely to occur.

TABLE 16.1 Unconditioned Motivating Operations (UMOs) That Increase Reinforcer Effectiveness and Evoke Relevant Behavior

UMO	Reinforcer-Establishing Effect	Evocative Effect
Food deprivation	Increases effectiveness of food as a reinforcer	Increases current frequency of all behavior previously reinforced with food
Water deprivation	Increases effectiveness of water as a reinforcer	Increases current frequency of all behavior previously reinforced with water
Sleep deprivation	Increases effectiveness of sleep as a reinforcer	Increases current frequency of all behavior previously reinforced with being able to sleep
Activity deprivation	Increases effectiveness of activity as a reinforcer	Increases current frequency of all behavior previously reinforced with activity
Oxygen deprivation[a]	Increases effectiveness of breathing (oxygen) as a reinforcer	Increases current frequency of all behavior previously reinforced with being able to breathe
Sex deprivation	Increases effectiveness of sex stimulation as a reinforcer	Increases current frequency of all behavior previously reinforced with sexual stimulation
Becoming too warm	Increases effectiveness of temperature decrease as a reinforcer	Increases current frequency of all behavior previously reinforced with becoming cooler
Becoming too cold	Increases effectiveness of temperature increase as a reinforcer	Increases current frequency of all behavior previously reinforced with becoming warmer
Increase in painful stimulus	Increases effectiveness of pain decrease as a reinforcer	Increases current frequency of all behavior previously reinforced with a decrease in painful stimulation

[a]It is not actually oxygen deprivation that functions as a UMO, but rather the buildup of carbon dioxide in the blood as a result of not being able to excrete carbon dioxide because of not being able to breathe or because of breathing in air that is as rich in carbon dioxide as the exhaled air.

TABLE 16.2 Unconditioned Motivating Operations (UMOs) That Decrease Reinforcer Effectiveness and Abate Relevant Behavior

UMO	Reinforcer-Abolishing Effect	Abative Effect
Food ingestion (after food deprivation)	Decreases effectiveness of food as a reinforcer	Decreases current frequency of all behavior previously reinforced with food
Water ingestion (after water deprivation)	Decreases effectiveness of water as a reinforcer	Decreases current frequency of all behavior previously reinforced with water
Sleeping (after sleep deprivation)	Decreases effectiveness of sleep as a reinforcer	Decreases current frequency of all behavior previously reinforced with sleep
Being active (after activity deprivation)	Decreases effectiveness of activity as a reinforcer	Decreases current frequency of all behavior previously reinforced with activity
Breathing (after not being able to breathe)	Decreases effectiveness of breathing (oxygen) as a reinforcer	Decreases current frequency of all behavior previously reinforced with being able to breathe
Orgasm or sexual stimulation (after sex deprivation)	Decreases effectiveness of sexual stimulation as a reinforcer	Decreases current frequency of all behavior previously reinforced with sexual stimulation
Becoming cooler (after being too warm)	Decreases effectiveness of temperature decrease as a reinforcer	Decreases current frequency of all behavior previously reinforced with becoming cooler
Becoming warmer (after being too cold)	Decreases effectiveness of temperature increase as a reinforcer	Decreases current frequency of all behavior previously reinforced with becoming warmer
Painful stimulation decrease (while in pain)	Decreases effectiveness of pain decrease as a reinforcer	Decreases current frequency of all behavior previously reinforced with a decrease in pain

In addition to establishing pain reduction as a reinforcer and evoking the behavior that has produced pain reduction, painful stimulation may evoke aggressive behavior toward another organism.[9] Painful stimulation may function as a UMO that establishes events such as signs of damage as effective reinforcers and evokes behaviors that have been reinforced by such signs. Skinner (1953) made this case in his analysis of anger and extended it to the emotions of love and fear.[10]

Although it is indisputable that both food deprivation/satiation and painful stimulation/reduction have unlearned value-altering effects, restriction of and access to certain conditioned reinforcers also seem to affect their value. This is the case with social reinforcers, such as praise (Vollmer & Iwata, 1991). It could be argued that for infants, attention (e.g., touch, parent's voices) is a form of unconditioned reinforcer whose value may be affected by deprivation and satiation. However, for most adults, attention functions as a conditioned reinforcer due to its relation to other forms of effective reinforcers (Michael, 2000). In this case, access and restriction to attention (or any other conditioned reinforcer) would function as a type of conditioned MO, as discussed below.

MOs FOR PUNISHMENT

MOs may also alter (increase or decrease) the punishing effectiveness of a stimulus, as well as the frequency (evoke or abate) of behaviors that have been punished by that stimulus. Thus, an EO related to punishment would have a punisher-establishing effect and an abative effect. For example, a migraine would establish the effectiveness of bright light as a punisher and decrease the frequency of all behaviors that in the past have produced the bright light (e.g., opening the curtains, turning a light on). Some biological states, such as pregnancy, may establish certain tastes or smells as punishers and decrease the occurrence of behaviors that have produced them. An AO related to punishment would have a punisher-abolishing effect and an evocative effect, in that behaviors that have been punished would occur again (assuming that a reinforcer contingency for the previously punished behavior is intact). Ingesting alcohol may decrease the effectiveness of social disapproval as a punisher and increase behaviors that have produced some social embarrassment (assuming those behaviors also produced some reinforcement in the past).

Most punishers that affect humans are effective because of a learning history; that is, they are *conditioned* rather than *unconditioned* punishers. In the case of conditioned punishers that were established by having been paired with the removal (or reduced availability) of reinforcers (i.e., negative punishment), their effectiveness as conditioned punishers will depend on the same MOs that establish the value of the reinforcers that were removed. For instance, a parent's disapproving look may function as a conditioned punisher as a result of having been paired with restricted access to video games. The disapproving look will be most effective as a conditioned punisher when an MO for playing video games is in effect. If the child has already played enough video games for the day (i.e., is "satiated"), then the disapproving look will be less effective in reducing behavior. Thus, *social disapproval* (e.g., a frown, a head shake, or vocal response such as "No" or "Bad") that typically precedes withholding of a reinforcer will only function as an effective conditioned punisher when the MOs for the withheld reinforcers are in effect. However, it is possible for a stimulus to be associated with several different punishers. In this case, this stimulus may function as a generalized conditioned punisher, whose effectiveness will be almost free from the influence of MOs. If the disapproving look has been associated with the loss of a wide range of reinforcers, including generalized reinforcers such as money, then the look

may almost always function as a punisher, since at least one of the reinforcers will be valuable at any given moment.

The same applies for the negative punishment procedures *time-out* and *response cost* (see Chapter 15). Time-out will only be effective when "time-in" serves as a reinforcer when the behavior is being punished. Solnick, Rincover, and Peterson (1977) compared the effectiveness of time-out when the time-in environment was either impoverished or enriched when trying to decrease self-injury and stereotypy in a child diagnosed with intellectual disabilities. The impoverished time-in condition looked no different from the child's typical treatment session, in which correct responses (sorting blocks) were reinforced with praise, food, and toys. During the enriched time-in condition, the experimenters played music, introduced new toys, and increased social interaction by continuously prompting the participant to play with the toys. Time-out consisted of 90 sec of no access to reinforcers contingent upon problem behavior. Their results showed that time-out was effective when the time-in was enriched and generally ineffective when time-in was impoverished. This suggests that the effectiveness of a negative punishment procedure depends upon the MO for the reinforcer being eliminated. The same can be said about *response cost*; unless the events removed contingent upon behavior are effective as reinforcers at the time the procedure is implemented, response cost will not serve as an effective punishment procedure.

In general, when observing a punishment effect, one must also consider the status of the reinforcer responsible for the occurrence of the punished behavior. A behavior is targeted for reduction only if it has been occurring and its occurrence is a function of reinforcement. For instance, disruptive behavior maintained by attention may increase in frequency when it has not received much attention (MO for reinforcement). Any punishment procedure targeting this disruptive behavior will have to be delivered when the MO for attention is in effect.

Suppose a time-out procedure had been used to punish disruptive behavior. As mentioned earlier, only when the MO relevant to the reinforcer available (attention during time-in) was in effect should we expect time-out to function as punishment. And then only if those MOs were in effect should we expect to see the abative effect of the punishment procedure on disruptive behavior (i.e., immediate reduction in behavior). In the absence of the MO for disruptive behavior (lack of attention), there would be no behavior to be abated.

Although these complex behavioral relations have not received much attention in the conceptual, experimental, or applied behavior analytic literatures, they seem to follow naturally from existing knowledge of reinforcement, punishment, and motivating operations. Behavior analysts should be aware of the possible participation of these behavioral relations in any situation involving punishment.[11]

MULTIPLE EFFECTS OF MOs

Most environmental events have more than one effect upon behavior. It is important to recognize these various effects and to not confuse one effect with another (Skinner, 1953, pp. 204–224). Multiple effects are readily apparent in an animal laboratory demonstration of an operant chain. A food-deprived rat pulls a cord hanging from the ceiling of the chamber. Pulling the cord produces an auditory stimulus such as a buzzer. In the presence of the buzzer, the rat presses a lever and receives a food pellet. The onset of the buzzer has two obvious effects: (1) It evokes the lever press, and (2) it increases the future frequency of cord pulling. The buzzer functions as an S^D for the evocative effect and as a conditioned reinforcer for the reinforcement effect. Both effects influence behavior in the same direction: an increase in the current frequency and an increase in future frequency of the lever press.[12] Similarly, a hungry person may open the fridge to find her roommate's leftover Mexican food. The sight of the Mexican food evokes behaviors associated with eating it (e.g., getting it out of the fridge, heating it up), as well as possibly increasing the future frequency of opening the fridge.

Similarly, if pressing the lever in the presence of the buzzer produces an electric shock, then the sound of the buzzer would function as a *discriminative stimulus for punishment* (S^{Dp}). The buzzer would then have abative effects on the current frequency of lever pressing and function as a conditioned punisher that decreases the future frequency of cord pulling. Likewise, if the leftover food is spoiled, then it abates the behavior of eating it as well as decreases the behavior of opening the fridge to look for food.

Like S^Ds, events that function as UMOs will typically have behavior-altering effects on the current frequency of a type of behavior and, when presented as consequences, will have function-altering effects on the future frequency of behavior that immediately preceded their onset. An increase in painful stimulation will, as an antecedent, have MO establishing effects that would *increase* the *current* frequency of all behavior that has alleviated pain and, as a consequence, would *decrease* the *future* frequency of whatever behavior immediately preceded the increased pain.

As in the example above, events that have a UMO evocative effect, such as painful stimulation, will also function as punishment for the response immediately preceding the onset of the event. However, this statement must be qualified for events that have such gradual onsets (like food or water deprivation) that they cannot easily function as consequences. In other words, even though food deprivation can evoke all behaviors that have produced food in the past, it would likely not be effective as a punishing consequence to decrease the future frequency of behavior.

Events that have MO abolishing effects as antecedents may have reinforcing effects when presented as consequences for behavior. For example, food consumption momentarily decreases the reinforcing value of food as an effective reinforcer (abolishing effect), as well as momentarily decreases behaviors that have produced food in the past (abative effect). As a consequence, food consumption increases the future frequency of behaviors that have produced it (reinforcing effect). Thus, it is important to note that when edibles are used as consequences for increasing behavior (reinforcement effect), the repeated consumption of edibles decreases their value as an effective reinforcer (abative effect). Similarly, if an item is removed in an attempt to increase its value as a reinforcer (establishing effect), its removal may reduce the future frequency of whatever behavior immediately

preceded the removal (punishing effect). Tables 16.3 and 16.4 illustrate the multiple effects of environmental events.

CONDITIONED MOTIVATING OPERATIONS (CMOs)

Motivating variables that alter the reinforcing effectiveness of other stimuli, objects, or events as a result of the organism's learning history are called **conditioned motivating operations (CMOs)**. As with UMOs, CMOs alter the momentary frequency of all behavior that has been reinforced by those other events. In commonsense terms, some environmental variables, as a product of our experiences, make us want something different from what we wanted prior to encountering those variables, and encourage us to try to obtain what we now want. For example, needing to enter a room with a locked door establishes the key to the lock as an effective reinforcer. A locked door is a CMO because its value-altering effect is a function of a learning history involving doors, locks, and keys.

There seem to be at least three kinds of CMOs, all of which are motivationally neutral events prior to their relation to another MO or to a form of reinforcement or punishment.

TABLE 16.3 Contrasting the Behavior-altering (MO Evocative) and Function-altering (Punishing) Effects of Environmental Events

Environmental Event	Evocative Effect on Current Behavior as a UMO	Function-altering Effect on Future Behavior as Punishment
Deprivation of food, water, sleep, activity, or sex	Increases the current frequency of all behavior that has been reinforced with food, water, sleep, activity, or sex	Should be punishment, but onset is too gradual to function as a behavioral consequence
Oxygen deprivation	Increases the current frequency of all behavior that has been reinforced with being able to breathe	The sudden inability to breathe decreases the future frequency of the type of behavior that preceded that instance of being unable to breathe
Becoming too cold	Increases the current frequency of all behavior that has been reinforced by becoming warmer	Decreases the future frequency of the type of behavior that preceded that instance of becoming too cold
Becoming too warm	Increases the current frequency of all behavior that has been reinforced by becoming cooler	Decreases the future frequency of the type of behavior that preceded that instance of becoming too warm
Increase in painful stimulation	Increases the current frequency of all behavior that has been reinforced with pain reduction	Decreases the future frequency of the type of behavior that preceded that instance of painful stimulation

TABLE 16.4 Contrasting the Behavior-altering (MO Abative) and Function-altering (Reinforcing) Effects of Environmental Events

Environmental Event	Abative Effect on Current Behavior as a UMO	Function-altering Effect on Future Behavior as Reinforcement
Ingestion of food and water, sleeping, being active, or having sex	Decreases the current frequency of all behavior that has been reinforced with food, water, sleep, activity, or sex	Increases the future frequency of the type of behavior that immediately preceded the food or water ingestion, being able to sleep, being able to be active, or sexual activity
Breathing (after not being able to breathe)	Decreases the current frequency of all behavior that has been reinforced with (oxygen) being able to breathe	The sudden ability to breathe increases the future frequency of the type of behavior that preceded that instance of ability to breathe
Becoming cooler (after being too warm)	Decreases the current frequency of all behavior that has been reinforced by becoming cooler	Increases the future frequency of the type of behavior that preceded that instance of becoming cooler
Becoming warmer (after being too cold)	Decreases the current frequency of all behavior that has been reinforced by becoming warmer	Increases the future frequency of the type of behavior that preceded that instance of becoming warmer
Decrease in painful stimulation	Decreases the current frequency of all behavior that has been reinforced with pain reduction	Increases the future frequency of the type of behavior that preceded that instance of pain decrease

The distinction among these conditioned motivating operations is based upon how they were developed and how they affect behavior. A **surrogate conditioned motivating operation (CMO)** produces the same effects as another MO because of some type of pairing (i.e., a surrogate is a substitute or stand-in), a **reflexive CMO** alters a relation to itself (makes its own removal effective as reinforcement), and a **transitive CMO** makes something else effective as reinforcement (rather than altering itself).

Surrogate CMO (CMO-S)

A CMO-S is a previously neutral stimulus that acquired its MO effects by being paired with a UMO. This process is similar to how respondent conditioned stimuli, operant conditioned reinforcers, and operant conditioned punishers acquire their functions by being paired with another behaviorally effective stimulus.

For example, a restaurant may exert discriminative control over several behaviors, including ordering food, since restaurants have been differentially correlated with the availability of food. However, because we typically go to restaurants after a period in which we have not had any food (i.e., food deprivation as a UMO), it is possible that after multiple pairings with food deprivation, the restaurant would function as a CMO-S, (a) establishing food as a reinforcer and (b) evoking behaviors that have produced food (e.g., ordering some). Figure 16.6 exemplifies the establishment of the restaurant as a CMO-S.

However, the existence of this type of CMO seems to be opposed to the organism's best interest for survival (Mineka, 1975). For example, ordering food when satiated, as in the example above, would not seem healthy. However, a number of animal studies have shown that a stimulus paired with a reinforcer (e.g., food) under a high (e.g., 24 h) deprivation condition is preferred over a stimulus that was paired with the same reinforcer amount under a low (e.g., 12 h) deprivation condition (e.g., Lewon & Hayes, 2014; Vasconcelos & Urcuioli, 2008). In other words, when two stimuli are paired with the same reinforcer, but under different MO conditions, subjects that are mildly deprived chose the stimulus correlated with high-deprivation levels. It is possible that in addition to the food deprivation (mild deprivation), the presence of the stimulus

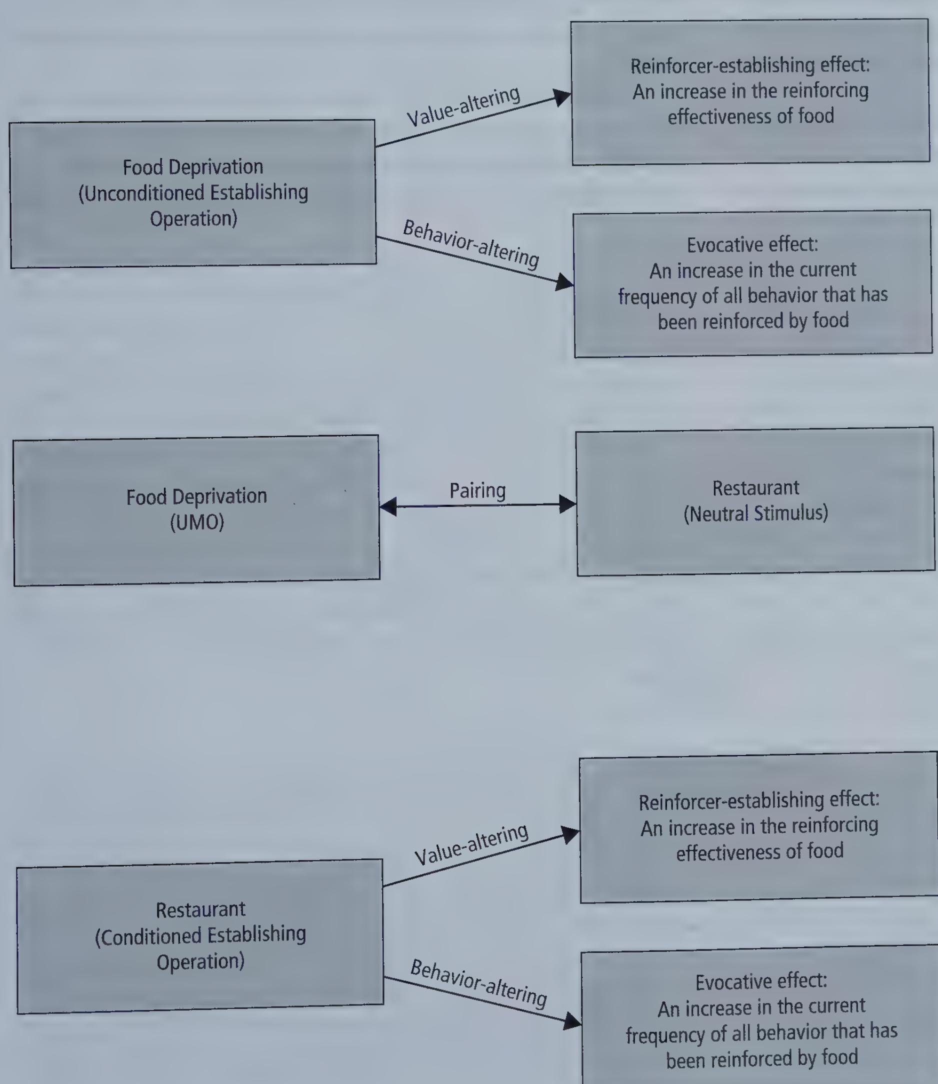


Figure 16.6 Example of a surrogate CMO (CMO-S).

correlated with high deprivation acts as a CMO-S, which makes the subjects behave as if they are "hungrier" than they would be otherwise, and this might be why they tend to respond more frequently to the high-deprivation alternative. Durlach, Elliman, and Rogers (2002) have found that adults under the same MO condition chose a flavored and colored drink that had been correlated with a high-value UMO for liquids (high-salt meal) over a drink that had been paired with a low-value UMO (low-salt meal), suggesting that the flavor and color acquired CMO-S properties.

O'Reilly, Lancioli, King, Lally, and Dhomhnaill (2000) found that aberrant behavior (pushing, pinching, property destruction, and self-injury) displayed by two individuals with developmental disabilities was maintained by attention when their parents were interacting with a third person (diverted attention). Given that parents provided low levels of attention when they were talking to a third person, diverted attention was paired with low attention (MO). This may have led to diverted attention functioning as a CMO-S. Figure 16.7 shows that for both participants, problem behavior that never occurred during the brief functional analysis did so when the diverted attention condition (CMO-S) was introduced. Providing attention noncontingently (abolishing operation) produced significant reductions in problem behavior. Follow-up observations showed reduced levels of problem behavior. These results seem to suggest that previously neutral stimuli may acquire MO functions due to being paired with another effective MO.

In a translational evaluation of the CMO-S, Lanovaz, Rapp, Long, Richling, and Carroll (2014) repeatedly paired a poster board with the presence of a condition (last 2 minutes of each session) known to function as an EO for stereotypy. After this pairing procedure, three out of four participants diagnosed with autism displayed higher rates of stereotypy in the presence of the poster board. These results have important applied implications, as there may be many environmental features in a therapeutic setting that could acquire CMO-S control over problem behavior simply by being paired with another MO. If a toy known for evoking stereotypy is delivered after compliance for a specific instruction, this repeated pairing may establish the instruction as a CMO-S that would increase the value of a specific form of automatic reinforcement and evoke stereotypy (Lanovaz et al., 2014).

The relationship between the CMO-S and an effective MO can generally be weakened by two kinds of **MO unpairing**: presenting the previously neutral stimulus (now CMO-S) without the effective MO or presenting the MO as often in the absence of the CMO-S as in its presence. In the examples above, if the instruction is no longer followed by the toy that usually evokes stereotypy, or a parent talking to a third person is no longer followed by low levels of attention, the value-altering and behavior-altering effects of the toy or diverted parental attention would be weakened. Similarly, if the toy known to evoke stereotypy is presented as often in the presence of the instruction as in its absence, or if differing levels of attention are associated with talking to a third person, their motivating effectiveness would be reduced.

The CMO-S appears to play a role in the development of behaviors that do not seem to make much sense, such as seeing a restaurant after dinner and wanting to eat again or feeling the urge to urinate after seeing the sign for the restroom. Both the

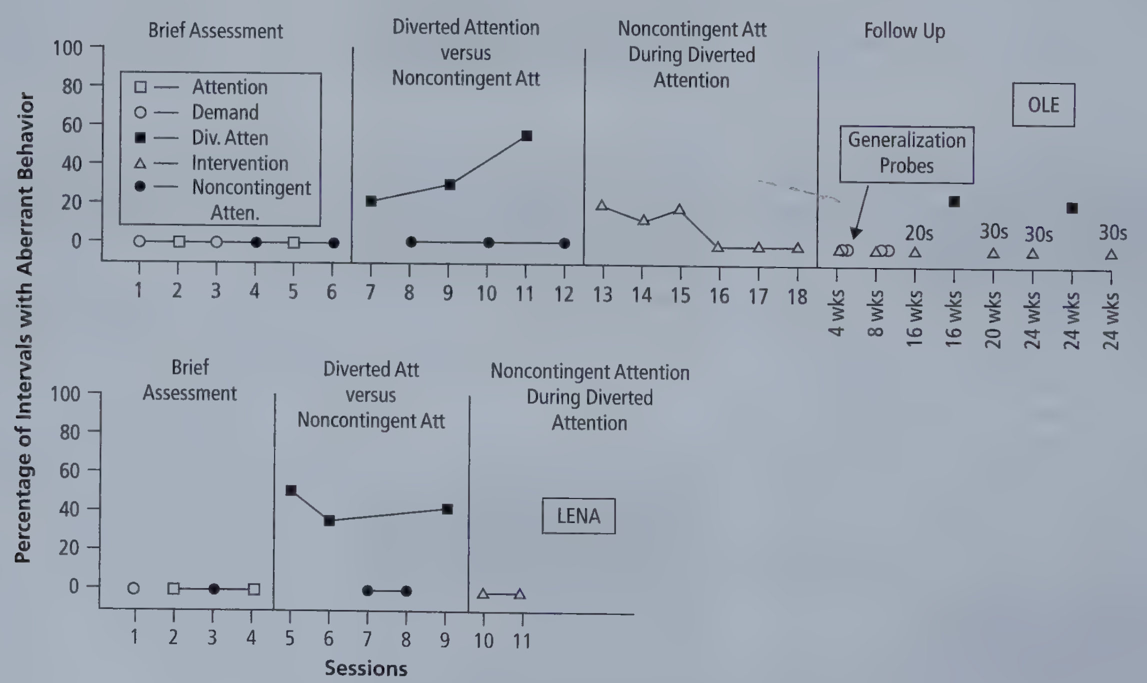


Figure 1. Percentage of intervals with aberrant behavior for Ole and Lena during assessment, treatment evaluation, and follow-up sessions.

Figure 16.7 Example of control by a motivating operation "a surrogate CMO (CMO-S)".

From "Using Brief Assessment to Evaluate Aberrant Behavior Maintained by Attention", M. F. O'Reilly, G. E. Lancioni, L. King, G. Lally, and O. N. Dhomhnaill, 2000, Reproduced with permission of John Wiley & Sons Inc.

restaurant and the restroom may evoke behavior as a CMO-S, given previous pairings with food deprivation and a full bladder, respectively (Miguel, 2013).

Reflexive CMO (CMO-R)

Any stimulus that systematically precedes the onset of painful stimulation becomes a CMO-R, in that its own offset functions as a reinforcer, and its occurrence evokes any behavior that has produced such reinforcement in the past. For example, in a traditional discriminated avoidance procedure,[13] after some interval of time, the onset of an initially neutral stimulus is followed by the onset of painful stimulation—usually electric shock (see Box 16.3, "What's an Aversive Stimulus?"). Some arbitrary response (one that is not part of the organism's phylogenic repertoire), such as lever pressing, terminates the painful stimulation (the animal *escapes* the pain) and restarts the interval. If the same response occurs during the warning stimulus, it terminates that stimulus and the shock does not occur on that trial. The response at this phase of the procedure is said to have avoided the pain and is called an *avoidance response*. Many organisms learn to respond during the onset of a warning stimulus, and thus receive very few shocks.

In this case, the warning stimulus functions similarly to the shock as an MO for the escape response, the reinforcer for which is shock termination. However, the capacity for the warning stimulus to establish its own termination as an effective reinforcer is a result of the organism's learning history, involving the pairing of the warning stimulus and shock. In other words, the warning stimulus evokes the avoidance response as a CMO, just as the painful stimulation evokes the escape response as a UMO.

The warning stimulus is not an S^D correlated with the *availability* of the consequence, but rather an MO related to the *reinforcing effectiveness* of the consequence. Recall that a discriminative stimulus is related to the current availability of a type of consequence for a given type of behavior. Availability is defined by two components: (a) An effective consequence (i.e., one whose MO is currently in effect) must have followed the response in the presence of the stimulus; and (b) the response must have occurred without the consequence (which would have been effective as a reinforcer if it had been obtained) in the absence of the stimulus. The relation between the warning stimulus and consequence availability does not meet the second requirement. In the absence of the warning stimulus, there is no effective consequence that could have failed to follow the response as an analog to the extinction responding that occurs in the absence of an S^D. The fact that the avoidance response does not turn off the absent warning stimulus is like the unavailability of food reinforcement for a food-satiated organism.

By contrast, a stimulus that has preceded some form of improvement may function as a CMO-R by (a) establishing its own offset as an effective punisher and (b) abating any behavior that has been so punished. For example, a positive evaluation from an employer may have been followed by a promotion or some sort of monetary bonus. This history would establish the offset of positive evaluations as a form of punishment and likely reduce the frequency of all behaviors that may have resulted in a bad evaluation (e.g., being late). Although this relation is quite plausible, little research has seemed to focus on this type of CMO-R.

The CMO-R plays an important role in identifying the negative aspects of many everyday interactions. Imagine that a stranger asks you where a particular building is located or asks for the time. The appropriate response is to give the information quickly or to say you don't know. Typically, the person who asked will smile and say, "thank you". Also, your question may be reinforced by the knowledge that a fellow human being has been helped (Skinner, 1957). In a sense, the question is an opportunity to obtain reinforcers that were not previously available. However, the question also begins a brief period that can be considered a warning stimulus, and if a response is not made soon, a form of social worsening (awkwardness) will occur. The stranger may repeat the question, stating it more clearly or more loudly, and will certainly think you are socially inept if you do not respond quickly. You would also consider it inappropriate to provide no answer. Even when no clear threat for nonresponding is implied by the person who asked, our social history under such conditions implies a form of worsening for continued inappropriate behavior. Many such situations probably involve a mixture of positive and negative components, but in those cases in which answering the question is an inconvenience (e.g., the listener is in a hurry), the stranger's thanks is not a strong reinforcer, nor is helping one's fellow human being, and the CMO-R is probably the main controlling variable.

Another example is when a person thanks another for doing a kindness of some sort. Saying "thanks" is evoked by the person's performing the favor or the kindness. Performing the favor may be considered an S^D in the presence of which one has learned to say "Thank you" and received reinforcement

BOX 16.3

What's an Aversive Stimulus?

A stimulus may be called aversive because its contingent removal functions as reinforcement, because its contingent presentation functions as punishment, because it evokes behavior that in the past has terminated it as an MO, or because it has evocative effects with respect to certain smooth muscle and glandular responses (e.g., heart rate increase, adrenal secretion) (Michael, 1993). Thus, *aversive stimulus* does not imply a specific functional relation; it is simply a behavioral translation of common sense expressions for "unpleasant feeling," "unpleasant states of mind," and so on—a form of usage possibly fostered by the term's lack of specificity.

consisting of the other person saying, "You're welcome." However, in many cases, ordinary courteous remarks may involve a CMO-R component. Consider the following scenario. Bill has his arms full, carrying something out of the building to his car. As he approaches the outer door, Danielle opens the door and holds it for Bill. Ordinarily, Bill would then smile and say, "Thanks." The CMO-R component can be illustrated by supposing that Bill just walks out without acknowledging the favor. In such circumstances, it would not be unusual for Danielle to call out sarcastically, "You're welcome!" Someone's favor may function as a warning stimulus (a CMO-R) that has systematically preceded some form of disapproval in the absence of some form of acknowledgement (saying "Thanks").

In the typical laboratory avoidance procedure, the response terminates the warning stimulus. In extending this analysis to the human situation, it must be recognized that the warning stimulus is not simply the event that initiated the interaction. In the previous example, the CMO-R is not the vocal request itself, which is too brief to be terminated. It is, instead, the more complex social stimulus situation consisting of having been asked and not having made a response during the time when such a response would be appropriate. The termination of that situation is the reinforcement for the response. Some social interactions with a stimulus—a facial expression, an aggressive posture—are more like the warning stimulus seen in the laboratory that can be terminated by the avoidance response, but most involve the more complex stimulus condition consisting of the request and the subsequent period described earlier.

In applied behavior analysis, the CMO-R is a component of training procedures for teaching individuals with defective social and verbal repertoires. During early intensive behavioral intervention, for example, learners are typically asked questions or given verbal instructions, followed by verbal or physical prompts so they can respond correctly. This arrangement may function as CMO-Rs, which will be followed by further intense social interaction if they do not respond appropriately (e.g., error correction). Such questions and instructions may be functioning primarily as CMO-Rs, rather than S^Ds related to the possibility of receiving praise or other positive reinforcers (Carbone et al., 2010). In other words, in these situations, it is likely that correct responses are partially maintained by negative rather than positive reinforcement. In addition to establishing their own removal as a reinforcer and evoking escape-maintained behaviors, these instructions may establish the value of activities or items unrelated to the task at hand as conditioned reinforcers. In this case, the instructions would function as transitive CMOs, as described in the following section (Michael, 2000). Figure 16.8 illustrates how, after preceding an aversive event, an instruction may become a CMO-R that would (a) establish its own removal as a conditioned reinforcer and (b) evoke any behavior that has produced such a reinforcer, such as compliance with the instruction.

Behaviors maintained by negative reinforcement are under the influence of a variety of antecedent stimuli, including task difficulty as a CMO-R. Changing these tasks may abolish the reinforcing effects of negative reinforcement and abate undesirable escape behaviors (Smith & Iwata, 1997). As an attempt to assess the function of problem behavior, as well as identify the specific features of the environment that served as EOs, McComas, Hoch, Paone, and El-Roy (2000) exposed three children diagnosed with developmental disabilities to a typical experimental functional analysis, a descriptive assessment to generate hypotheses about possible antecedents for destructive behavior, and an analysis in which they tested the evocative effects of the hypothesized EOs. Although all participants engaged in problem behavior in the presence of task demands, the specific task features that evoked problem behavior (CMO-Rs) varied across participants. However, problem behavior by all participants was reduced simply by making specific modifications to the instructional methods, without modifying the demands. For one participant (Eli), problem behavior was evoked by a novel or difficult task (CMO-R). During the EO analysis, the authors compared the level of problem behavior when an instructional strategy was used (i.e., manipulables available during a math task) versus when the strategy was not used. Problem behavior rarely occurred when the strategy was used, suggesting that the establishing and evocative effects of task demands on escape-maintained problem behavior were eliminated by the addition of the instructional strategy. This effect may have been achieved by making the tasks less difficult and, thus, unpairing task demands with aversive properties.

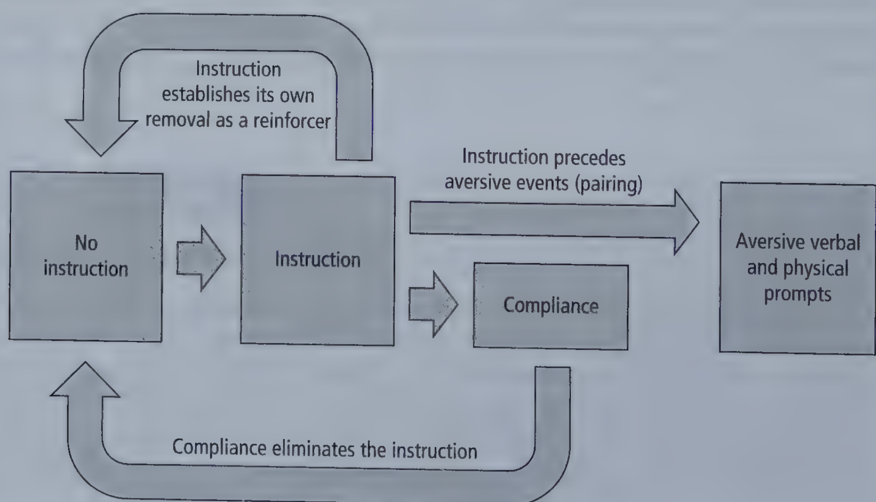


Figure 16.8 Illustration of how an instruction may become a reflexive conditioned motivating operation (CMO-R).

Results from a sample participant are shown in Figure 16.9. Functional analysis (top panel) suggested that problem behavior consisted of escape from academic tasks. During the EO analysis and follow-up (middle panel), problem behavior did not occur when the instructional strategy was employed, compared to when it was not. Additionally, the participant was more compliant when the instructional strategy was available (bottom panel).

Numerous treatments to weaken or abolish CMO-Rs that evoke problem behavior in the presence of task demands have been evaluated (see Carbone et al., 2010, for a review). These include programming competing reinforcers for compliance (e.g., Piazza et al., 1997), pairing stimuli that evoke problem behavior with reinforcement by embedding preferred activities during task presentation (e.g., Kemp & Carr, 1995), reducing the aversiveness of the task by using errorless procedures (Heckaman, Alber, Hooper, & Heward, 1998), systematically fading instructional demands (Pace, Iwata, Cowdery, Andree, & McIntyre, 1993), varying the tasks being presented (McComas et al. 2000; Winterling, Dunlap, & O'Neill, 1987), increasing the pace of instruction to produce higher rates of reinforcement for appropriate behavior (Roxburgh & Carbone, 2013), presenting highly preferred activities and low demands to neutralize CMO-Rs occurring prior to the session (e.g., Horner, Day, & Day, 1997), reducing the aversiveness of tasks by providing choices of activities and reinforcers during sessions (Dyer, Dunlap, & Winterling, 1990), interspersing easy tasks with difficult ones (Mace & Belfiore, 1990), gradually introducing novel tasks (Smith, Iwata, Goh, & Shore, 1995), and modifying task duration based on whether participants engage in problem behavior early or late in the session (Smith et al., 1995). It may also be possible to weaken the CMO-R by unpairing it with the aversive stimulus it precedes (Kettering, Neef, Kelley, & Heward, 2018).

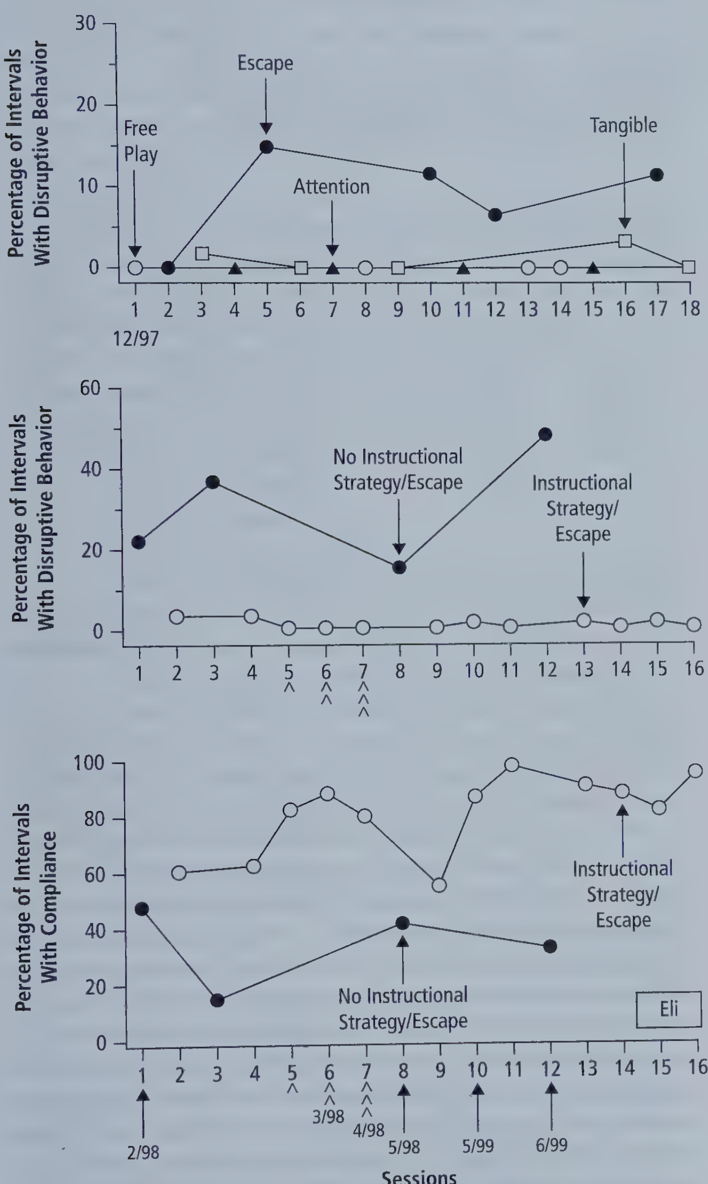


Figure 1. Percentage of intervals with shirt biting for Eli in the functional analysis (top panel) and EO analysis and follow-up (middle panel). Percentage of intervals with compliance is depicted in the bottom panel. Follow-up sessions begin in Session 5. The single carat indicates probe in novel setting, the double carat indicates probe with novel instructor, and the triple carat indicates probe with novel materials.

Figure 16.9 Example of control by a motivating operation a reflexive CMO (CMO-R).

From "Escape Behavior During Academic Tasks: A Preliminary Analysis of Idiosyncratic Establishing Operations" by J. J. McComas, H. Hoch, H., D. Paone, and D. El-Roy 2000, Reproduced with permission of John Wiley & Sons Inc.

It is important to note that negatively reinforced problem behaviors improve environmental conditions that are otherwise poor. Thus, it is the behavior analyst's ethical duty to identify and modify these aversive variables (including CMO-Rs), rather than simply reduce problem behavior through some form of consequence manipulation such as escape extinction (Carbone et al., 2010; McGill, 1999).

Transitive CMO (CMO-T)

When an environmental variable establishes the effectiveness of another event as a reinforcer or punisher, it is referred to as a CMO-T. This type of MO was previously described in the behavior analytic literature as an *establishing stimulus* (S^E; Michael, 1982) or *blocked-response CEO* (Michael, 1988; also see Miguel, 2013). Conditioned reinforcers are dependent upon a variable functioning as a CMO-T. Thus, all variables that function as UMOs also function as CMO-Ts for the stimuli that are conditioned reinforcers because of their relation to the relevant unconditioned reinforcer. For example, not only is food deprivation a UMO that momentarily (a) establishes the value of food as an unconditioned reinforcer and (b) evokes all behaviors that have produced food in the past; food deprivation is also a CMO-T that momentarily (a) establishes the value of all conditioned reinforcers paired with food (e.g., an attentive server in a restaurant, a menu, utensils) and (b) evokes all behaviors that have produced these conditioned reinforcers in the past.[14]

Consider the simple operant chain described earlier: A food-deprived rat pulls a cord that turns on a buzzer. In the presence of the buzzer sound, the rat then presses a lever and receives a food pellet. Food deprivation as a UMO makes food effective as an unconditioned reinforcer, a relation that requires no learning history. Food deprivation as a CMO-T also makes the buzzer sound effective as a conditioned reinforcer, which

clearly requires a learning history. Thus, food deprivation is a UMO with respect to the reinforcing effectiveness of food, and a CMO-T with respect to the reinforcing effectiveness of the buzzer sound. Conversely, food ingestion is a UMO that abolishes the reinforcing effectiveness of food and the buzzer sound.

In the Mexican restaurant, we may need to call the waiter to get his attention prior to ordering our tacos. In this scenario, food deprivation functions as a UMO that establishes food as an unconditioned reinforcer, and as a CMO-T that establishes the sight of the waiter as a conditioned reinforcer. As in the rat example above, food ingestion would decrease the reinforcing values of both food and the waiter.

Other UMOs, such as painful stimulation, also establish the reinforcing value of conditioned reinforcers. As mentioned above, as a UMO, a headache would establish its own removal as an unconditioned reinforcer. However, as a CMO-T, a headache also establishes the reinforcing value of other stimuli (e.g., aspirin) that have served to reduce headaches.

The relationship between UMOs and conditioned reinforcers cannot be overlooked by clinicians or educators who may make extensive use of conditioned reinforcers, including generalized ones, to strengthen academic or other socially significant behaviors. Studies have shown that the value of generalized conditioned reinforcers in the form of tokens decreases when participants have access to the primary and conditioned reinforcers paired with the tokens (Ivy, Neef, Meindl, & Miller, 2016; Moher, Gould, Hegg, & Mahoney, 2008). Conversely, when participants have restricted access to back-up reinforcers, tokens are more effective reinforcers, as evidenced by an increased rate of responding. In practice, these data suggest that the value of conditioned and generalized conditioned reinforcers (e.g., tokens) depends upon the MO relevant to their back-up reinforcers. Thus, it seems reasonable to assume that the effectiveness of generalized reinforcers would be increased when paired with a greater number of back-up reinforcers (Skinner, 1957, p. 54). More specifically, satiation may be prevented by pairing the putative generalized reinforcer with a variety of back-up reinforcers, such as food and drink that are under control of dissimilar MOs (Moher et al., 2008). Money, for instance, seems almost completely free from MO control, given that it has been paired with a multitude of effective reinforcers related to several different MOs. Thus, it is likely that one or more of these MOs will be in effect when money is delivered contingent on behavior, making its effectiveness almost impossible to abolish.

The reinforcing effectiveness of many (probably most) conditioned reinforcers is not only altered by relevant UMOs, as described above, but also dependent on other stimulus conditions. This is why the effectiveness of conditioned reinforcers is often said to be "context" dependent. When the context is not appropriate, the stimuli may be available but are not accessed because they are not effective as reinforcers. A change to an appropriate context will evoke behavior that has been followed by those stimuli, which are now effective as conditioned reinforcers. The occurrence of the behavior is not related to the availability, but rather to the value, of its consequence. For example, flashlights are usually available in home settings but are not accessed until a power failure makes them valuable. In this sense, the power failure (the sudden darkness) *evokes* the behavior that has obtained the flashlight in the past (rummaging around in a particular drawer). However, when the light is back on, the value of the flashlight as a reinforcer is abolished, and any behaviors that led to the flashlight are momentarily abated. The motivating nature of this CMO-T relation is not widely recognized, and the evocative variable (the sudden darkness) in this example is usually interpreted as an S^D.

Consider a workman disassembling a piece of equipment, with his assistant handing him tools as he requests them.[15] The workman encounters a slotted screw that must be removed and requests a screwdriver. The sight of the screw evoked the request, the reinforcement for which is receiving the tool. Prior to an analysis in terms of the CMO-T, the sight of the screw would have been considered an S^D for the request, but such screws have *not* been differentially related to the availability of reinforcement for requests for tools. In the typical workman's history, assistants have provided requested tools irrespective of the stimulus conditions in which the request occurred. The sight of the screw is more accurately interpreted as a CMO-T for the request, not as an S^D.

The fact that several S^Ds are involved in this complex situation makes the analysis somewhat difficult. The screw *is* an S^D for unscrewing movements (with a screwdriver in hand) or for selecting a screwdriver rather than another tool. The workman's verbal request for a screwdriver, although evoked by the sight of the slotted screw as a CMO-T, is dependent on the assistant's presence as an S^D. The offered screwdriver is also an S^D for the workman reaching to grasp it. The critical issue here, however, is the screw's role in evoking the request, and this is a motivating rather than a discriminative relation (Michael, 1982).

Another common example involves a stimulus related to some form of danger that evokes some relevant protective behavior. Imagine a night security guard is patrolling an area and hears a suspicious sound. He pushes a button on his phone that signals another security guard, who then activates his phone and asks if help is needed (which reinforces the first guard's call). The suspicious sound is not an S^D in the presence of which the second security guard's response is more available (the second guard would respond to the first guard's call whether or not the first guard heard the sound), but rather a CMO-T in the presence of which second guard's response is more valuable. S^Ds are involved, however. A ringing phone *is* an S^D in the presence of which one has activated the phone, said something into the receiver, and been reinforced by hearing a response from another person. Answering phones that are not ringing has typically not been so reinforced.[16] (Note, incidentally, that the effect of the danger signal is not to evoke behavior that produces its own termination, as with the CMO-R, but rather behavior that produces some other event, in this case, the sound of the security guard's colleague offering to help.)

Figure 16.10 shows an example of sudden darkness functioning as a CMO-T that establishes a flashlight as a conditioned reinforcer, as well as evoking the behavior of looking for a flashlight. Note that sudden darkness may also function as a CMO-R that establishes its own removal as a reinforcer.[17]

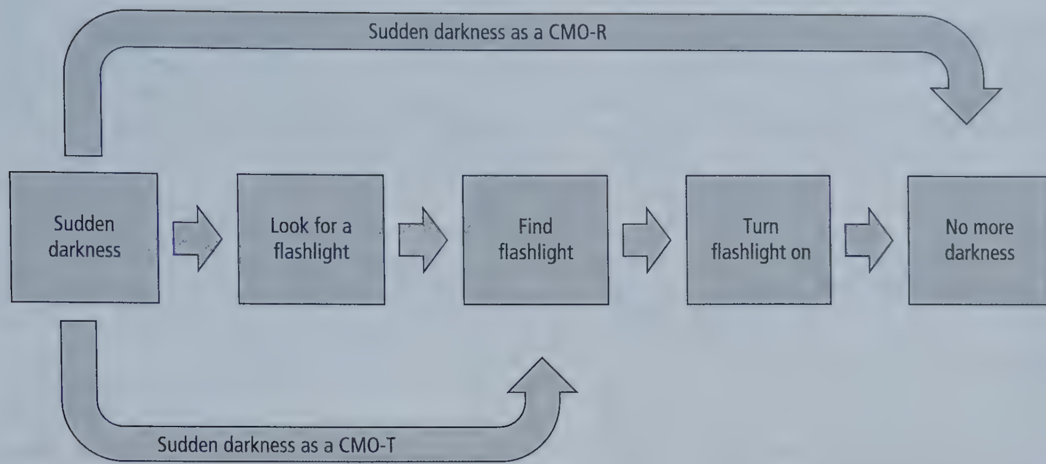


Figure 16.10 Illustration of a transitive conditioned motivating operation (CMO-T).

The distinction between an S^D and a CMO-T hinges on the relation between reinforcer availability and the presence or absence of the stimulus. If the reinforcer is more available in the presence than in the absence of the stimulus, the stimulus is an S^D; if the reinforcer is equally available in the absence and presence of the stimulus, the stimulus is a CMO-T. Screwdrivers have typically been available to the workman in the absence as well as in the presence of screws. The response by the security guard's colleague has been just as available in the absence as in the presence of a suspicious noise.

A CMO-T evokes behavior because of its relation to the *value* of a consequence rather than to the *availability* of a consequence. The two forms of behavioral control, the S^D and the CMO-T, which are so different in origin, would be expected to differ in other important ways. This is an example of terminological refinement in behavior analysis, not a discovery of new empirical relations. The value of this refinement will be found in the improved theoretical and practical effectiveness of behavior analysts whose verbal behavior has been affected by it.

Teaching Mands with CMO-Ts. Mand training is an essential aspect of language programs for individuals with severely deficient verbal repertoires (see Chapter 18). For such individuals, manding does not spontaneously arise from tact and receptive discrimination training (i.e., listener behavior). The learner has to want something, make an appropriate verbal response, and be reinforced by receiving what was wanted. This procedure places the response under control of the relevant MO. The manipulation of CMO-Ts allows the teacher to contrive a situation to get the learner to want something that can be a means to another end (e.g., obtaining information by asking a question). Any stimulus, object, or event can be a basis for a mand simply by arranging an environment in which that stimulus can function as a conditioned reinforcer. Thus, if a pencil mark on a piece of paper is required for an opportunity to play with a favorite toy, a mand for a pencil and for a piece of paper can be taught.

The manipulation of CMO-Ts is common practice when teaching mands to individuals with disabilities (e.g., Alwell, Hunt, Goetz, & Sailor, 1989; Betz, Higbee, & Pollard, 2010; Goetz, Gee, & Sailor, 1985; Lechago, Carr, Grow, Love, & Almason, 2010; Shillingsburg, Valentino, Bowen, Bradley, & Zavatkay, 2011). The typical procedure, termed *interrupted-chain* (Hall & Sundberg, 1987) involves withholding an item the learner needs to complete a previously learned task. This manipulation functions as a CMO-T, increasing the value of the item as a reinforcer, which creates the ideal condition for teaching the learner to mand for it. In a mand training study, Shillingsburg, Bowen, Valentino, and Pierce (2014) exposed three children with disabilities to two scenarios. One scenario involved varying the placement of a preferred item in one of nine different colored (or animal) cups. A stimulus that signaled the preferred item (e.g., an empty Skittles bag) was then placed next to the participant. When the participant asked for the item (e.g., "May I have a Skittle?"), the experimenter responded by saying that the item was under one of the cups, increasing the value of the information for the location of the item (which cup) as a conditioned reinforcer. The second scenario involved giving a highly preferred item to one of three adults when the participant was not looking. The stimulus that signaled the presence of the preferred item was presented again, and when the child asked for it, the therapist said that one of the other adults had it, without saying whom. This statement served as a CMO-T that increased the value of the information about who had the item as a conditioned reinforcer. The authors used a vocal prompt delay to teach participants to ask, "which cup?" and "who has it?" in scenarios 1 and 2, respectively. These CMO-T trials were alternated with trials in which the information was not valued as a reinforcer (abolishing operation; AO) to guarantee that behavior came under control of the EO, versus some other aspect of the environment. During AO trials, the therapist provided either the information about the location of the item or the name of the person who had it. In addition to the emission of mands, the authors also collected data on whether participants selected the correct cup or approached the correct adult to obtain the preferred item. All participants manded using "which" and "who" more often during EO rather than AO conditions, and approached the preferred items during both conditions, suggesting that the interrupted chain procedure was successful in placing mands under functional control of CMO-Ts. Figure 16.11 shows the results for the "which" questions. None of the participants manded for information during

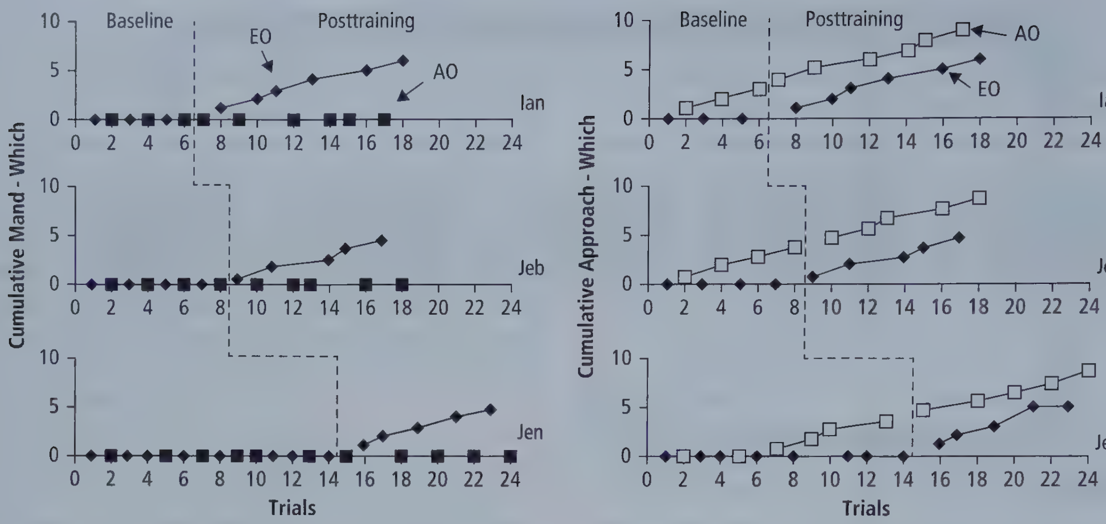


Figure 1. Cumulative record of independent responses for manding "which?" (left) and cumulative record for approach behavior (right) during baseline and posttraining probes for Ian, Jeb, and Jen.

Figure 16.11 Example of control by a motivating operation a transitive CMO (CMO-T).

From "Mands for Information Using "Who?" and "Which?" in the Presence of Establishing and Abolishing Operations" by M, A. Shillingsburg, C. N. Bowen, A. L. Valentino, and L. W. Pierce, 2014, Reproduced with permission of John Wiley & Sons Inc.

baseline (left panel), although each selected the correct container during the AO condition during which the information on the location of the item was provided (right panel). After training, participants manded only during the EO condition (left panel). Additionally, after receiving the information during the EO condition, participants started to select the correct containers to access the preferred item (right panel). Similar results were obtained for the "who" questions. Chapter 18 provides additional details about using CMO-Ts during mand training.

RELEVANCE OF MOs TO THE GENERALITY OF TREATMENT EFFECTS

In the applied area, the reinforcer-establishing effect of MOs seems to be well understood, as clinicians may temporarily withhold edibles, tangibles, and even social attention to make them more effective as reinforcers for behaviors being taught (e.g., Taylor et al., 2005). However, these target behaviors will not occur in future or novel circumstances, even if well learned and part of the learner's repertoire, unless relevant MOs are in effect (Fragale et al., 2012; O'Reilly et al., 2012). Fragale and colleagues have shown that after teaching three children diagnosed with autism to mand for preferred items and testing for generalization across settings, the newly acquired mands occurred almost exclusively when items had been restricted prior to the session and seldom occurred when participants could access the items prior to sessions. These results show that even if well established, a behavior may not occur in the future in a new setting if the MO is absent. Thus, it is possible that generalization and maintenance failures are not only a function of novel (generalization) stimuli being dissimilar to those present during training, but also due to different MOs. Thus, when programming for generalization, clinicians must do so not only across different discriminative stimuli (Stokes & Baer, 1977) but also across different MOs (Miguel, 2017). In summary, for a newly acquired response to be generalized and maintained, its relevant MO must also be in effect.

RELEVANCE OF MOs TO APPLIED BEHAVIOR ANALYSIS

Applied behavior analytic research has shown the effects of MOs on a multitude of behaviors across a variety of conditions. MOs have affected the results of preference assessments (Gottschalk, Libby, & Graff, 2000; McAdam et al., 2005) and functional analyses (Fischer, Iwata, & Worsdell, 1997; Worsdell, Iwata, Conners, Kahng, & Thompson, 2000), engagement with activities (Klatt et al., 2000; Vollmer & Iwata, 1991), within-session responding (Roane, Lerman, Kelley, & Van Camp, 1999), social initiations (Taylor et al., 2005), the rate of acquisition of target behaviors (Ivy et al., 2016), and the occurrence of challenging behaviors (O'Reilly et al., 2007).

Gottschalk and colleagues (2000), for instance, compared the results of preference assessments of four children with autism under three conditions. In the control condition, participants had access to four edible items in premeasured portions, 3 times, 24 hr prior to the assessment. In the satiation condition, the same access to premeasured portions was provided, but the participants had free access to one of the stimuli for 10 min prior to the assessment. Finally, in the deprivation condition, participants had the same premeasured portions, but were deprived of one stimulus for 48 hr prior to the assessment. Results showed that across all participants, approach responses were higher during deprivation conditions. These results were replicated with leisure items with both typically developing children and individuals with intellectual disabilities (McAdam et al., 2005) and suggest that results of preference assessments are highly dependent on current MOs.

As an attempt to evaluate the effects of MOs on the results of functional analyses, Worsdell et al. (2000) exposed six individuals with profound intellectual disabilities who engaged in self-injury to an experimental functional analysis (FA), followed by another FA in which the MO and the reinforcer for problem behavior were present or absent. During the EO condition, the experimenter ignored all but problem behavior, while in the AO condition, the experimenter delivered noncontingent reinforcement. These conditions were associated with either the presence or the absence of continuous reinforcement. The authors reported high rates of problem behavior for all participants during the condition in which the EO and reinforcement were present, and almost no responding when both the EO and reinforcement were absent, or when the EO was present and reinforcement was absent. These results are consistent with the conceptualization that for behavior to occur, it must be followed by an *effective* reinforcer, and suggest that MOs may differentially affect the results of FAs.

As previously discussed, in addition to affecting results of commonly used behavior analytic assessments, MOs have been shown to directly affect the rate of acquisition of novel behaviors reinforced with primary, conditioned, or generalized reinforcers (e.g., Ivy et al., 2016; Lang et al., 2009; Moher et al., 2008; O'Reilly et al., 2009), as well as the frequency of behaviors that have already been acquired (e.g., Klatt et al., 2000). Behavior analysts have also directly manipulated MOs to reduce problem behaviors (e.g., Rapp, Cook, McHugh, & Mann, 2017) and teach new skills (Lechago et al., 2010).

Research on MOs falls within the three broad categories previously described by Iwata, Smith, and Michael (1991): demonstrating the effects of MOs on behavior, using MOs to clarify results of assessments, and manipulating MOs to increase or decrease behavior. This ever-growing literature has had much impact on our practice. It allowed behavior analysts to recognize that the three-term contingency is dependent upon a fourth variable that modulates the effectiveness of reinforcement, and hence the evocative strength of discriminative stimuli. A clear understanding of these motivational variables seems critical for interpreting assessment results and designing effective behavior analytic interventions.

As we go about our daily lives, we may clearly identify how MOs influence our own behavior. Revisiting the anecdote about our obsession with Mexican food, we see that food deprivation as a UMO, or, more specifically, not having had Mexican food for some time, increases its value as a reinforcer, and the frequency of previous behaviors that have produced it, such as going to our favorite Mexican restaurant. Food deprivation as a CMO-T establishes the sight of the restaurant as a conditioned reinforcer, increasing any behavior that preceded it (e.g., driving there) and making us happy when we see it. When at the restaurant, the sight of tortilla chips may increase the value of salsa as a reinforcer, as they tend to covary (CMO-S). The spiciness of the salsa establishes its own removal as a reinforcer (CMO-R), which makes us drink water or any other liquid (e.g., too many margaritas). When the chips are gone (CMO-T), we call the waiter to bring us more. And even if we have already eaten, the sight of the restaurant alone (CMO-S) may get us to eat more. Given the ubiquity of motivational variables in our lives, a complete understanding of human behavior demands an analysis of the critical roles of MOs.

SUMMARY

Definition and Characteristics of Motivating Operations

1. A motivating operation (MO) (a) alters the effectiveness of some stimulus as a reinforcer, the value-altering effect; and (b) alters the current frequency of all behavior that has been reinforced by that stimulus, the behavior-altering effect.
2. The value-altering effect is either (a) an increase in the reinforcing effectiveness of some stimulus, in which case the MO is an establishing operation (EO); or (b) a decrease in reinforcing effectiveness, in which case the MO is an abolishing operation (AO).
3. The behavior-altering effect is either (a) an increase in the current frequency of behavior that has been reinforced by some stimulus, called an *evocative effect*; or (b) a decrease in the current frequency of behavior that has been reinforced by some stimulus, called an *abative effect*.
4. The value-altering effect of an MO changes the potency of reinforcers (or punishers) to increase and maintain (or decrease and eliminate) behavior
5. The behavior-altering effect is a change in frequency due to the organism's behavior contacting reinforcement or punishment. This change occurs only *after* the consequence has been obtained.
6. The value-altering and behavior-altering effects of an MO occur simultaneously but are considered independent in the sense that one does not derive from the other. However, once behavior contacts reinforcement, these two effects may not be discernible.
7. The behavior-altering effect of the MO can directly increase or decrease the frequency of behavior. It can also indirectly affect the evocative or abative strength of relevant discriminative stimuli.
8. The value-altering effect of the MO can directly alter the reinforcing (or punishing) effectiveness of the relevant consequence. It can also indirectly alter the reinforcing (or punishing) effectiveness of any (conditioned) stimulus correlated with the relevant consequence.
9. S^Ds and MOs alter the current frequency of behavior (behavior-altering effect).

10. Reinforcers, punishers, and response occurrence without consequence alter the future frequency of behavior (function-altering effect).

Distinguishing Between MOs and S^Ds

11. An S^D controls a type of behavior because it has been related to the *differential availability* of an effective reinforcer for that type of behavior. This means that the relevant consequence has been available in the presence of, and unavailable in the absence of, the stimulus. Most variables that qualify as motivating operations fail to meet this second S^D requirement because in the absence of the variable, there is no MO for the relevant reinforcer, and thus no reinforcer unavailability.

12. A useful contrast is that S^Ds are related to the differential availability of a currently effective form of reinforcement for a particular type of behavior, whereas MOs are related to the differential reinforcing effectiveness of a particular event.

13. In nontechnical terms, an S^D tells you that something you want is available; an MO makes you want something.

Unconditioned Motivating Operations (UMOs)

14. For all organisms there are events, operations, and stimulus conditions with unconditioned *value-altering* effects. In other words, these MOs establish or abolish the value of stimuli as reinforcers in the absence of prior learning.

15. The main UMOs for humans are those related to deprivation and satiation with respect to food, water, oxygen, activity, and sleep; and those related to sexual reinforcement, comfortable temperature conditions, and painful stimulation. For each variable, there are two MOs, one with an establishing operation (EO) and one with an abolishing operation (AO). Also, each variable has an evocative effect and an abative effect. Thus food deprivation is an EO and has evocative effects on relevant behavior, and food ingestion is an AO and has abative effects on relevant behavior.

MOs for Punishment

16. MOs alter (increase or decrease) the punishing effectiveness of a stimulus, as well as the frequency (evoke or abate) of behaviors that have been punished by that stimulus.

17. An EO related to punishment would have a punisher-establishing effect and an abative effect.

18. An AO related to punishment would have a punisher-abolishing effect and an evocative effect, in that behaviors that have been punished would occur again (assuming that a reinforcer contingency for the previously punished behavior is intact).

19. In the case of conditioned punishers that were established by having been paired with the removal (or reduced availability) of reinforcers (i.e., negative punishment), their effectiveness as conditioned punishers will depend on the same MOs that establish the value of the reinforcers that were removed.

20. Social disapproval, time-out from reinforcement, and response cost are stimulus conditions that usually function as punishment because they are related to a reduction in the availability of some kinds of reinforcers. The MOs for those forms of punishment are the MOs for the reinforcers that are being made less available.

Multiple Effects of MOs

21. Most environmental events have more than one effect upon behavior.

22. Events that function as UMOs will typically have behavior-altering effects on the current frequency of a type of behavior and, when presented as consequences, will have function-altering effects on the future frequency of behavior that immediately preceded their onset.

23. Events that have MO evocative effects, such as painful stimulation, will also function as punishment for the response immediately preceding the onset of the event.

24. Events that have MO abolishing effects, such as food consumption, may have reinforcing effects when presented as consequences for behavior.

Conditioned Motivating Operations (CMOs)

25. Motivating variables that alter the reinforcing effectiveness of other stimuli, objects, or events, but only as a result of the organism's learning history, are called *conditioned motivating operations* (CMOs). As with the UMOs, they also alter the momentary frequency of all behavior that has been reinforced (or punished) by those other events.

26. The surrogate CMO (CMO-S) is a stimulus that acquires its MO effectiveness by being paired with another MO, and has the same value-altering and behavior-altering effects as the MO with which it was paired.

27. The CMO-S appears to play a role in the development of behaviors that do not seem to make much sense, such as seeing a restaurant after dinner and wanting to eat again or feeling the urge to urinate after seeing the sign for the restroom. Both the restaurant and the restroom may evoke behavior as a CMO-S, given previous pairings with food deprivation and a full bladder, respectively.

28. A stimulus that acquires MO effectiveness by preceding some form of worsening or improvement is called a *reflexive CMO* (CMO-R). It is exemplified by the warning stimulus in a typical escape–avoidance procedure, which establishes its own offset as reinforcement and evokes all behavior that has accomplished that offset.

29. A stimulus that has preceded some form of improvement may also function as a CMO-R by (a) establishing its own offset as an effective punisher and (b) abating any behavior that has been so punished.

30. The CMO-R has usually been interpreted as an S^D. The CMO-R fails to qualify as an S^D, however, because in its absence there is no MO for a reinforcer that could be unavailable and thus no reinforcer unavailability.

31. The CMO-R identifies a negative aspect of many everyday interactions that otherwise would be interpreted as a sequence of opportunities for positive reinforcement. One example is a request for information, which initiates a brief period during which a response must be made to terminate a period of increased social awkwardness.

32. The CMO-R is often an unrecognized component of procedures used in teaching effective social and verbal behavior. Learners are asked questions or given instructions, which are followed by further intense social interaction if they do not respond to them appropriately. The question or instruction may be functioning more as a warning stimulus, as a CMO-R, than as an S^D related to an opportunity for receiving praise or other positive reinforcers.

33. An environmental variable that establishes (or abolishes) the reinforcing effectiveness of another stimulus and evokes (or abates) the behavior that has been reinforced by that other stimulus is a transitive CMO, or CMO-T.

34. Variables that function as UMOs also function as transitive CMOs for the stimuli that are conditioned reinforcers because of their relation to the unconditioned reinforcer. Food deprivation (as a UMO) establishes as a reinforcer not only food but also (as a CMO-T) all of the stimuli that have been related to obtaining food (e.g., the utensils with which one transports food to the mouth).

35. The reinforcing effectiveness of many conditioned reinforcers not only is altered by relevant UMOs but also may be dependent on other stimulus conditions. Those stimulus conditions, as CMO-Ts, then also evoke behavior that has obtained the conditioned reinforcers.

36. A CMO-T evokes behavior because of its relation to the *value* of a consequence rather than to the *availability* of a consequence.

37. The manipulation of CMO-Ts is common practice when teaching mands to individuals with disabilities. The typical procedure, termed *interrupted-chain*, involves withholding an item the learner needs to complete a previously learned task. This manipulation functions as a CMO-T, increasing the value of the item as a reinforcer, which creates the ideal condition for teaching the learner to mand for it.

Relevance of MOs to the Generality of Treatment Effects

38. An acquired behavior may not occur in the future in a new setting if the MO is absent. Thus, it is possible that generalization and maintenance failures are not only a function of novel (generalization) stimuli being dissimilar to those present during training, but also due to different MOs.

39. When programming for generalization, clinicians must do so not only across different discriminative stimuli but also across different MOs. For a newly acquired response to be generalized and maintained, its relevant MO must also be in effect.

Relevance of MOs to Applied Behavior Analysis

40. Applied behavior analytic research has shown the effects of MOs on a multitude of behaviors across a variety of conditions. MOs have affected the results of preference assessments and functional analyses, engagement with activities, within-session responding, social initiations, the rate of acquisition of target behaviors, and the occurrence of challenging behaviors.

41. Research on MOs allowed behavior analysts to recognize that the three-term contingency is dependent upon a fourth variable that modulates the effectiveness of reinforcement, and hence the evocative strength of the discriminative stimuli.

42. A clear understanding of MOs seems critical not only for interpreting and assessing results but also for designing effective behavior analytic interventions.

43. Given the ubiquity of motivational variables in our lives, a complete understanding of human behavior demands an analysis of the critical roles of MOs.

KEY TERMS

abative effect
abolishing operation (AO)
behavior-altering effect
conditioned motivating operation (CMO)
establishing operation (EO)
evocative effect
function-altering effect
motivating operation (MO)
reflexive conditioned motivating operation (CMO-R)
reinforcer-abolishing effect
reinforcer-establishing effect
surrogate conditioned motivating operation (CMO-S)
transitive conditioned motivating operation (CMO-T)
unconditioned motivating operation (UMO)
MO unpairing
value-altering effect

MULTIPLE-CHOICE QUESTIONS

1. _________ alters the effectiveness of some object or event as a reinforcer and the current frequency of all behavior that has been reinforced by that stimulus, object, or event.
   a. Discriminative stimulus
   b. Establishing operation
   c. Behavior-altering effects
   d. Value-altering effects

   Hint: (See "Definition and Characteristics of Motivating Operations")

2. The term _________ has been suggested to replace the term establishing operation (EO).
   a. Motivating operation
   b. Evocative effect
   c. Abative effect
   d. Abolishing operation

   Hint: (See "Definition and Characteristics of Motivating Operations")

3. The effect that can produce either an increase or decrease in the reinforcing effectiveness of some stimulus, object, or event is the:
   a. Behavior-altering effect
   b. Abative effect
   c. Evocative effect
   d. Value-altering effect

   Hint: (See "Definition and Characteristics of Motivating Operations")

4. The effect that can produce either an increase or decrease in the current frequency of behavior that has been reinforced by some stimulus, object, or event is the:
   a. Behavior-altering effect
   b. Abative effect
   c. Evocative effect
   d. Value-altering effect

   Hint: (See "Definition and Characteristics of Motivating Operations")

5. Motivating operations (MOs) and S^Ds are both
   a. Antecedent variables that alter the future frequency of some behavior
   b. Consequent variables that alter the current frequency of some behavior
   c. Antecedent variables that alter the current frequency of some behavior
   d. Consequent variables that alter the future frequency of some behavior

   Hint: (See "Distinguishing Between MOs and S^Ds")

6. _________ control(s) a type of behavior because it has been related to the differential availability or an effective reinforcer for that type of behavior.
   a. Motivating operations (MOs)
   b. Unconditioned motivating operations (UMOs)
   c. Conditioned motivating operations (CMOs)
   d. Discriminative stimulus (S^D)

   Hint: (See "Distinguishing Between MOs and S^Ds")

7. _________ is (are) related to the differential reinforcing effectiveness of a particular type of environmental event.
   a. Motivating operations
   b. Discriminative stimulus
   c. Motivating variables
   d. Discriminative operations

   Hint: (See "Distinguishing Between MOs and S^Ds")

8. This type of motivating operation has value-altering effects that are not learned.
   a. Conditioned motivating operations
   b. Unconditioned motivating operations
   c. Surrogate motivating operations
   d. Reflexive motivating operations

   Hint: (See "Unconditioned Motivating Operations (UMOs)")

9. Which of the following is an example of a UMO as it relates to humans?
   a. Getting paid for a job every two weeks
   b. Putting on a sweater to go outside
   c. Sleeping after a week of not getting any sleep
   d. Playing with your children after they have been in school all day

   Hint: (See "Unconditioned Motivating Operations (UMOs)")

10. The type of motivating operation that accomplishes what the original motivating operation it was paired with did is a
    a. Surrogate CMO
    b. Transitive CMO
    c. Reflexive CMO
    d. Substantive CMO

    Hint: (See "Conditioned Motivating Operations (CMOs)")

11. This type of motivating operation alters a relation to itself by acquiring MO effectiveness by preceding a worsening or improvement.
   a. Surrogate CMO
   b. Transitive CMO
   c. Reflexive CMO
   d. Substantive CMO
   Hint: (See "Conditioned Motivating Operations (CMOs)")

12. This type of motivating operation makes something else effective as a reinforcer because of its relation or association with an unconditioned reinforcer.
   a. Surrogate CMO
   b. Transitive CMO
   c. Reflexive CMO
   d. Substantive CMO
   Hint: (See "Conditioned Motivating Operations (CMOs)")

13. This type of motivating operation has particular implications for the training of language in individuals with little or no speech.
   a. Surrogate CMO
   b. Transitive CMO
   c. Reflexive CMO
   d. Substantive CMO
   Hint: (See "Conditioned Motivating Operations (CMOs)")

14. The continued study of motivating operations is most important to understanding in what field of applied behavior analysis?
   a. The three-term contingency
   b. How people learn
   c. Why people learn
   d. How to best refine the terminology
   Hint: (See "Relevance of MOs to Applied Behavior Analysis")

ESSAY-TYPE QUESTIONS

1. Define and discuss the characteristics of an establishing operation, or motivating operation (MO), and the recent movement to changing terminology.
   Hint: (See "Definition and Characteristics of Motivating Operations")
2. Discuss the similarities and differences between motivating operations (MOs) and discriminative relations.
   Hint: (See "Distinguishing Between MOs and S^Ds")
3. Define and discuss examples of unconditioned motivating operations (UMOs) related to the human organism.
   Hint: (See "Unconditioned Motivating Operations (UMOs)")
4. Given the unconditioned motivating operation (UMO) of food deprivation for an organism, describe the reinforcer-establishing and evocative effect that accompanies the UMO. Suppose food is subsequently ingested by the organism. Now describe the reinforcer-abolishing and abative effects.
   Hint: (See "Unconditioned Motivating Operations (UMOs)" and chapter examples)
5. Discuss the evocative effect on current behavior as a UMO and the function-altering effect on future behavior as punishment given the environmental variable of becoming too cold.
   Hint: (See "Unconditioned Motivating Operations (UMOs)" and chapter examples)
6. Define conditioned motivating operations (CMOs) and the various types of CMOs, and how each of these types operates.
   Hint: (See "Conditioned Motivating Operations (CMOs)")
7. What are some implications for the use of motivating operations (MOs) in the study and future of behavior analysis?
   Hint: (See "Relevance of MOs to Applied Behavior Analysis")

NOTES

1. Skinner (1938) used the term *drive* to express the relationships between operations other than reinforcement, such as deprivation, and changes in behavior. Even though Skinner found the term useful to describe these functional relations, he acknowledged it to be an inferred state between these operations and behavior and noted that the term was "not actually required in a descriptive system" (p. 368).
2. For an account of the history of the concept, see Miguel (2013).
3. The direct evocative effect suggests that the MO evokes behavior directly, in the absence of an S^D. However, we must be careful not to ignore that one of the defining features of an operant is the stimuli in the presence of which it occurs (Catania, 2013).
4. Michael (1986, 2004) used the terms "repertoire-altering effect" and "function-altering effect" to classify operations (e.g., reinforcement) that would change the future probability of an organism's behavior, as well as alter the function of stimuli.
5. Even though this is an example of manipulating an MO (i.e., task difficulty), the same argument applies to the manipulation of S^Ds. If problem behavior is more likely to occur in the presence of Dad (S^D) and not in the

presence of Mom (S^{Δ}), removing Dad from the room (eliminating the S^D) would only temporarily reduce (abate) the frequency of problem behavior. When Dad returns, the child is likely to misbehave again.

6. An individual's quality of life may be threatened by socially impoverished and/or extremely aversive environments, which may lead to the development of attention and escape-maintained problem behaviors. In addition to being an ethical necessity, improving the environment would reduce current problem behavior and help prevent its future occurrence.
7. It is true that pain reduction as a form of reinforcement is unavailable in the absence of pain, and also, food reinforcement is, in a sense, unavailable in the absence of food deprivation, but this is not the kind of unavailability that occurs in discrimination training and that develops the evocative effect of an S^D (Michael, 1982).
8. The terms *unconditioned* and *conditioned* modify MOs in the same way they modify respondent-eliciting stimuli and operant reinforcers and punishers. The effects of unconditioned motivating operations, like unconditioned eliciting stimuli for respondent behavior and unconditioned reinforcers and punishers, are not dependent on the organism's learning history. The effects of conditioned motivating operations, like the effects of conditioned eliciting stimuli and conditioned reinforcers and punishers, *are* dependent on a learning history.
9. Some of this aggression may be elicited by pain functioning as a respondent unconditioned stimulus (Ulrich & Azrin, 1962).
10. For a discussion of Skinner's approach to emotional predispositions in the general context of MOs, see Michael (1993, p. 197).
11. It is sometimes argued that behavioral principles are too simplistic for analyzing complex human behavior, and therefore some nonbehavioral—usually cognitive—approach is required. It may be that the behavioral repertoire of the people who make this argument is too simple for the task. However, behavioral principles themselves are not so simple, as can be seen from the preceding effort to understand the effects of punishment MOs, or the learned functions of painful stimulation.
12. The sound of the buzzer will also function as a respondent conditioned eliciting stimulus for smooth muscle and gland responses typically elicited by food in the mouth, and will condition such responses to any other stimulus that is present at that time—an instance of higher order classical conditioning (see Chapter 2)—but the current chapter discusses operant relations only.
13. The term *discriminated avoidance* arose so that this type of procedure could be distinguished from an avoidance procedure in which an exteroceptive stimulus change does not precede the shock itself (called *avoidance without a warning stimulus, non-discriminated avoidance*, or *free-operant avoidance*). See Chapter 12.
14. The evocative effect of CMO-Ts is not easily confused with the evocative effect of an S^D. If one can see food deprivation as an MO (rather than an S^D) with respect to the behavior that has been reinforced with food, then its function as an MO (rather than an S^D) with respect to behavior that has been reinforced with the various food-related conditioned reinforcers is an easy extension.
15. This scenario was first described in Michael (1982). At that time CMO-T was called an *establishing stimulus*, or S^E.
16. See Chapter 2, Box 2.2, "When the Phone Rings."
17. If removal of darkness is to be considered an unconditioned reinforcer, then sudden darkness would establish the value of its removal as a UMO.

CHAPTER 17

Stimulus Control

LEARNING OBJECTIVES

- Define and provide examples of stimuli and stimulus classes.
- Define and provide examples of stimulus control.
- Define and provide examples of establishing operations.
- Define and provide examples of stimulus generalization and stimulus discrimination.
- Explain how differential reinforcement is used to establish stimulus control (i.e., discrimination training procedures).
- Explain how stimulus control and stimulus generalization are used to produce concept formation.
- Describe how to use response and stimulus prompts to establish stimulus control.
- Describe how to use response and stimulus prompt fading to transfer stimulus control to the relevant stimulus.

In addition to selecting and maintaining operant behavior, reinforcement confers controlling function to antecedent stimuli whose presence or absence coincides with the organism's differential contact with it. Drivers step on the brakes when approaching stop signs, red traffic lights, and street-crossing pedestrians more often than they do in the absence of those stimuli. Bilingual speakers switch languages to match the language of their listeners. Baseball pitchers throw fastballs, curves, or change-ups, depending on signals from their catchers.

Control by antecedent stimuli plays a fundamental role in social, educational, treatment, and therapeutic environments. It can help us understand how behaviors considered appropriate in one context are inappropriate when emitted in another context. Loud talking is appropriate on the playground, but not in the classroom. Arriving 15 to 20 minutes late is appropriate for a party, but not for a job interview. Some behaviors considered inappropriate by parents, teachers, and society in general are not behavior problems per se. The problem is emitting behaviors at times, in places, or in circumstances deemed inappropriate by others. This represents a problem of stimulus control, and it is a major concern for applied behavior analysts.

This chapter describes basic concepts and processes for understanding stimulus control, methods for developing stimulus control, and techniques for transferring stimulus control from contrived antecedents to naturally occurring stimuli.

STIMULUS CONTROL: BASIC CONCEPTS AND PROCESSES

This section describes basic concepts and processes essential to understanding stimulus control: stimulus discrimination, stimulus generalization, faulty stimulus control, conditional stimulus control, respondent eliciting and operant discriminative functions of antecedent stimuli, and discriminative and motivational functions of antecedent stimuli.

Stimulus Discrimination

In a basic laboratory demonstration of operant conditioning, a rat is placed in an experimental chamber and given the opportunity to press a lever. Following each press the rat receives a food pellet, and the rate of lever pressing increases. This simple demonstration of reinforcement as a *two-term contingency* (operant response → reinforcement) seldom occurs outside of laboratory research (Pilgrim, 2015).

Manipulating the presence of antecedent stimuli in concert with the availability of reinforcement makes the contingency more complex and demonstrates how an organism's behavior can come under control of the environmental context. For example, a buzzer sounds occasionally and the rat receives a food pellet after lever presses when the buzzer is sounding, but lever presses in the absence of the buzzer sound produce no food pellets. Under these experimental conditions, the buzzer sound is called a **discriminative stimulus** (S^D, pronounced "ess-dee") and its absence is called **stimulus delta** (S^Δ, pronounced "ess-delta"). After some experience with this arrangement, the rat will press the lever more often in the presence of the buzzer sound than in its absence.

This *three-term contingency* (discriminative stimulus → *operant response → reinforcement*) is "considered the basic unit of analysis in the analysis of operant behavior" (Glenn, Ellis, & Greenspoon, 1992, p. 1332). Behavior that occurs more often in the presence of an S^D than in its absence is said to be under **stimulus control**. Technically, stimulus control occurs when the rate, latency, duration, or magnitude of a response is altered in the presence of an antecedent stimulus (Dinsmoor, 1995a, 1995b).

A stimulus acquires control only when responses emitted in the presence of that stimulus produce more reinforcement than do responses in the absence of stimulus. When the S^D signals a rich schedule of reinforcement and the S^Δ signals a zero chance of reinforcement (i.e., extinction), the organism will eventually respond only in the presence of the S^D and emit no responses in the S^Δ condition. In such cases, **stimulus discrimination** is achieved.

As introduced in Chapter 2, an S^D evokes behavior because its presence has been correlated with the *differential availability* of an effective reinforcer. That is, responses emitted in the presence of the S^D have produced reinforcement in the past, and responses emitted in its absence (S^Δ) have resulted in no reinforcement (extinction) or reinforcement of lesser quality or on a leaner schedule than the reinforcement available in the presence of the S^D (Michael, 2004). The development of stimulus control does not require an all-or-nothing face-off of reinforcement versus extinction conditions. While taking a break from working on this chapter by raking leaves, one of the authors of this text experienced an everyday example of stimulus control created by an S^Δ condition that entailed less reinforcement than the S^D condition.

> After raking several big piles of leaves from the front lawn to the curb, where the city vacuums them up for compost, I dropped the rake and took off my jacket. I picked up the rake and started on a new patch of lawn. But this time, each response with the same sweeping motion of the rake that was so efficient moments before moved fewer leaves. I looked down and saw I was holding the rake with its tines facing up (S^Δ) instead of down (S^D). From then on, I held the rake tines down!

Stimulus Generalization

In contrast to stimulus discrimination, **stimulus generalization** refers to the extent to which stimuli other than the S^D acquire stimulus control over the behavior. Stimuli sharing similar physical properties with the S^D are most likely to acquire evocative function. Stimulus generalization and discrimination are relative concepts. Stimulus generalization reflects a loose degree of stimulus control, whereas stimulus discrimination indicates a tight degree of control. Stimulus generalization can be observed when a young child who has been reinforced for saying "daddy" in the presence of her father says "daddy" in the presence of her father's two brothers. Further conditioning in the form of differential reinforcement—praise or other reinforcers for "daddy" in the presence of the child's father or photos and other likenesses of him and no reinforcement for "daddy" in the presence of her two uncles—will sharpen the degree of stimulus control to one stimulus class, the child's father and likenesses of him.

Stimulus generalization occurs with new stimuli that share similar physical dimensions with the controlling antecedent stimulus. For instance, if a behavior has a history of producing reinforcement in the presence of a blue stimulus, stimulus generalization is more likely with a lighter or darker color of blue than with a red or yellow stimulus. Also, stimulus generalization is likely when the new stimuli have other elements (e.g., size, shape) in common with the controlling stimulus. A student whose behavior has produced reinforcement for making a response to a circle is more likely to make the same response to an oval shape than to a triangular shape.

A **stimulus generalization gradient** graphically depicts the degree of stimulus generalization and discrimination by showing the extent to which responses reinforced in one stimulus condition are emitted in the presence of untrained stimuli. The slope of the gradient indicates the degree of generalization. When the slope of the gradient is relatively flat, little stimulus control is evident. However, an increasing slope of the gradient shows more stimulus control. The gradients are typically presented as rate, count (i.e., total number), or percentage of total number across all stimuli (see Figure 17.1).

Behavior analysts have used several procedures to produce stimulus generalization gradients. The classic technique of Guttman and Kalish (1956) provides a representative example. Their technique is important because many prior researchers had obtained stimulus generalization gradients by conditioning groups of subjects on the same stimulus value and then testing them individually, each with a different stimulus value. Obviously, this type of technique cannot demonstrate the degree of stimulus control for individual subjects. Guttman and Kalish provided a method of acquiring gradients for each subject and laid the foundation for greater understanding of the principles governing stimulus control.

Guttman and Kalish (1956) reinforced pigeons on a variable interval (VI) 1-min schedule for pecking a disk illuminated with a light source appearing yellow-green to humans (i.e., a wavelength of 550 mμ). After the rate of pecking the yellow-green disk had stabilized, the pigeons were tested under extinction conditions on the original stimulus and a randomized series of 11 different wavelengths never presented during training as the test for stimulus generalization.

Stimulus generalization occurs with responses to a new stimulus after a response has been conditioned in the presence of another similar stimulus. If responding produces reinforcement during testing for stimulus generalization, it cannot be clear

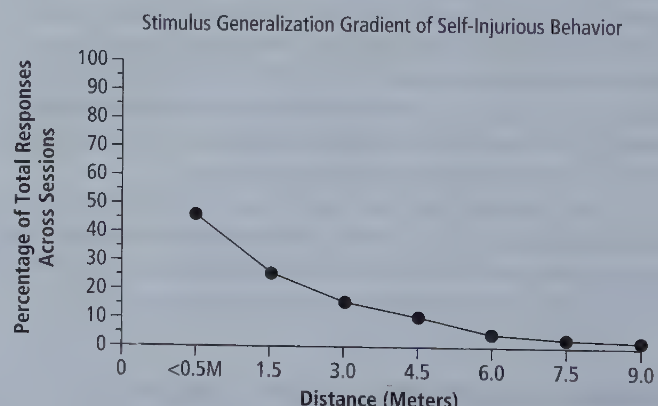


Figure 17.1 Percentage of total responses across sessions at a given distance during generalization tests. <0.5, 1.5, 3.0, 4.5, 6.0, 7.5, and 9.0 refer to the distance (in meters) between the therapist and the participant.

"Assessment of Stimulus Generalization Gradients in the Treatment of Self-Injurious Behavior", J. S. Lalli, F. C. Mace, K. Livezey, and K. Kates, 1998, Journal of Applied Behavior Analysis, 31, p. 481. Reproduced with permission of John Wiley & Sons Inc.

whether any responses to a new stimulus after the first response represent generalization or if the responses are a function of the reinforcement schedule. Guttman and Kalish (1956) avoided this problem of confounding their results by testing for generalization during extinction.

Lalli, Mace, Livezey, and Kates (1998) reported an excellent applied example of assessing and displaying a stimulus generalization gradient. They used a stimulus generalization gradient in an assessment of the relation between the physical proximity of an adult and the self-injurious behavior (SIB) of a 10-year-old girl with severe intellectual disabilities. The results presented in Figure 17.1 show that the percentage of total SIB across sessions became progressively lower as the distance between the therapist and the girl increased.

Faulty Stimulus Control

Faulty stimulus control occurs when a behavior comes under the restricted control of an irrelevant antecedent stimulus. Instructional materials students can complete with 100% accuracy without attending to the critical stimulus features of the material are widely used in primary and secondary classrooms. Figure 17.2 is an example of a worksheet likely to engender faulty stimulus control. Students need not read anything but the highlighted word at the end of each sentence to determine which picture goes with it. Other examples include a science vocabulary worksheet that can be completed without reading the definitions—the student simply matches the number of letters in each term with the number of spaces provided next to the definitions, and a language arts activity in which students construct compound words by drawing lines between the color-matched boxes (e.g., *base* and *ball* are in blue boxes, *bath* and *tub* in green boxes) without reading the words and thinking about which ones go together (Heward, 2003). Because students' answers on such poorly designed materials are under the stimulus control of irrelevant features (e.g., color, size, position on the page), the materials provide no meaningful practice with the knowledge or skills they were intended to teach.

Vargas (1984) suggested that a "Yes" answer to any of the following questions indicates that an instructional material is likely to result in faulty stimulus control over students' behavior:

1. Can students use pictures or diagrams instead of text to complete an exercise?
2. Does highlighting or physical layout give away answers, making it unnecessary to read through an assignment?

1. This is very **slick.**
2. Jim will go on the hill with his **sled.**
3. Lucy has a big hole in her **sleeve.**
4. It is time for me to go to **sleep.**
5. The dog has my **slipper.**

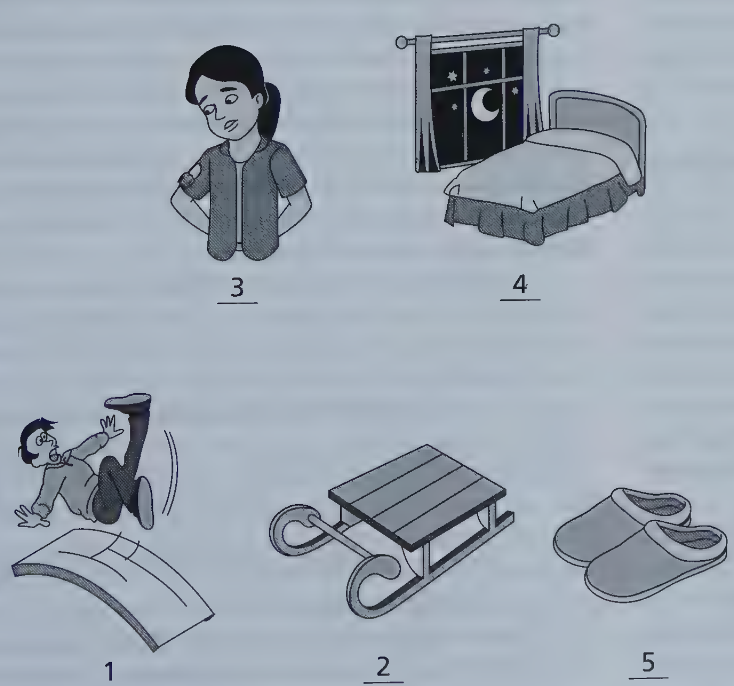


Figure 17.2 Example of instructional materials likely to engender faulty stimulus control.

From "What Are Your Exercises Teaching: An Analysis of Stimulus Control in Instructional Materials," by J. S. Vargas, 1984. In W. L. Heward, T. E. Heron, D. S. Hill, & J. Trap Porter (Eds.), *Focus on Behavior Analysis in Education* (p. 132). Copyright: Charles E. Merrill Publishing Company.

3. Are students able to answer questions on a passage without reading it?
4. Do all of the problems on a page require the same process for solution, making it unnecessary to discriminate between strategies?
5. Are the questions Jabberwocky comprehension questions—that is, can they be answered using grammatical cues alone? (p. 130)

Applied behavior analysts who work in schools should seek to detect and eliminate faulty stimulus control in instructional materials as their roles permit.

Conditional Stimulus Control

In a simple stimulus discrimination, only one antecedent stimulus controls the response. Holding a cup under the tap on a water cooler and pushing the blue lever is a simple discrimination by a person wanting cold water (see Figure 11.3). The blue lever is an S^D that signals the availability of cold water. A child touching a circle from an array of different shapes when her teacher says, "Touch the circle" is a simple discrimination. Most stimulus control in everyday life is more complex than the simple discriminations described by the three-term contingency: $S^D \rightarrow R \rightarrow S^{R+}$. Obtaining reinforcement often depends on discriminating the environmental context provided by multiple stimuli.

In a **conditional discrimination**, the response that will produce reinforcement in the presence of a particular stimulus depends on (is conditional on) the presence or absence of other stimuli. Pushing the blue (or red) lever when a friend asks, "Cold (or hot) water please," is a conditional discrimination. The friend's request is the conditional stimulus, and the color of the water tap lever is the discriminative stimulus. Conditional discrimination is evidenced in a child's correct response to the request "Point to the circle that is the same color as your shirt." The color of the child's shirt is the conditional stimulus that makes the same-colored circle the discriminative stimulus.

Respondent Eliciting and Operant Discriminative Functions of Antecedent Stimuli

At first glance, stimulus control of operant behavior appears similar to the control of respondent behavior by a conditioned stimulus. Discriminative stimuli and conditioned stimuli are antecedent stimuli whose presence increases the frequency of some behavior. Applied behavior analysts, however, need to distinguish between the evocative function of S^Ds for operant behavior and the eliciting function of conditioned stimuli for respondent behavior. This is a crucial distinction for understanding antecedent control of operant behavior.

In a typical laboratory demonstration of respondent conditioning, the experimenter presents food to a dog. The food is an unconditioned stimulus that elicits the unconditioned response—salivation. The experimenter then introduces a buzzer sound (a neutral stimulus). The buzzer sound does not elicit salivation. Then after several trials of pairing the buzzer sound with food delivery, the buzzer becomes a conditioned stimulus that elicits salivation (a conditioned response) in the absence of food (an unconditioned stimulus).

Laboratory experiments on operant and respondent conditioning have demonstrated consistently that antecedent stimuli can acquire control over a behavior. A buzzer sounds, and a rat presses a bar. A buzzer sounds, and a dog salivates. Despite the similarities, the bar press is an operant behavior, salivation is a respondent behavior, and the manner in which the S^D and the conditioned stimulus acquire their controlling functions is very different. An S^D acquires its controlling function for operant behavior by being paired with stimulus changes that occur immediately following behavior. Conversely, a conditioned stimulus acquires its controlling function for respondent behavior by being paired with other antecedent stimuli that elicit the behavior (i.e., an unconditioned stimulus or a conditioned stimulus).

Discriminative and Motivational Functions of Antecedent Stimuli

Distinguishing the nature of antecedent control of operant behavior can be difficult. Discriminative stimuli and establishing operations share two important similarities: Both occur before the behavior of interest and both have evocative functions. To evoke a behavior means to occasion it or call it up. The question often becomes, "Was a particular response evoked by a discriminative stimulus (S^D), an establishing operation (EO), or both?"

In some situations, an antecedent stimulus change evokes a response and the functional relation appears to be an S^D. For example, in a typical shock–escape procedure an animal is placed in an experimental chamber. Shock is administered until a response terminates the shock for a designated period. Then the shock is reintroduced until it is again terminated by a response, and so on. An experienced animal escapes the shock immediately. In such a situation, some would say that the shock serves as an S^D. The shock, an antecedent stimulus, evokes a response that is negatively reinforced (the shock is removed). In this situation, however, shock does not function as an S^D. To repeat an important point, an S^D evokes behavior because its presence has been correlated with the *differential availability* of an effective reinforcer. That is, responses emitted in the presence of the S^D have produced reinforcement in the past and responses emitted in its absence (S^Δ) have not.

Even though the animal receives reinforcement by terminating the shock, the absence of shock does not signal the availability of less frequent reinforcement. Before the response can be reinforced, the shock must be on. In this example, the onset of shock is an EO because it changes what *functions* as reinforcement (shock termination) rather than signaling the differential availability of reinforcement (Michael, 2000). Often the apparent S^D effect does not have a history of effective differential reinforcement correlated with the altered frequency of response. These situations are probably related to motivating operations (MOs) rather than stimulus control.

The following scenario puts the laboratory examples into an applied context: A teacher asks a student to begin working on an academic task. The student refuses to comply, disrupting the class, and the teacher withdraws the request. Later, the teacher

again requests an academic response, and the cycle of noncompliance and disruptive behavior and removal of the request continues. Although the teacher's request is an antecedent stimulus that evokes aggression that is negatively reinforced (the request is removed), it is not an S^D. Like the onset of shock in Michael's laboratory example, the teacher's request is an EO that evokes the aggressive behaviors rather than an S^D. It makes no sense to talk about an S^D evoking the aggressive behavior in the absence of the request (McGill, 1999), just as it makes no sense to talk about an S^D evoking a response in the absence of the shock. The laboratory rat is not motivated to escape shock in the absence of shock, and the student does not "want" to escape his teacher's task requests in the absence of those task requests. An antecedent stimulus functions as an S^D only when in its presence a specific behavior produces reinforcement, and in the absence of that stimulus the same behavior does not produce reinforcement (Michael, 2000).

The laboratory and classroom examples can be changed to show the difference between the evocative functions of EOs and stimulus control. The experimental conditions could change so that a buzzer sounds at different periods throughout the session, the shock would be removed only when a response is made while the buzzer is sounding, and a response would not produce reinforcement in the absence of the buzzer (i.e., the shock would not be removed). Under these conditions the buzzer would function as an S^D, and stimulus control would be demonstrated.

The classroom conditions could change so that two teachers work with the student. One teacher terminates task requests when the student is noncompliant; the other teacher does not remove the requests. In this arrangement, the presence of the first teacher signals the availability of negative reinforcement; the second teacher signals extinction. The teacher who terminates the requests to work would become an S^D that evokes noncompliant, disruptive behaviors from the student; the second teacher would become an S^Δ for noncompliance. In these modified examples, the characteristics of the antecedent control would be different in the presence and absence of the buzzer and the other teacher, and the buzzer and the other teacher would be correlated with an increased frequency of reinforcement.

An understanding of the S^D and MO evocative functions will improve technological descriptions of antecedent control and the understanding of behavior change (Laraway, Snycerski, Michael, & Poling, 2001). Ultimately, these understandings will produce greater effectiveness in education and treatment (see also Chapter 16).

DEVELOPING STIMULUS CONTROL

This section describes the major processes required for the development of stimulus control.

Stimulus Discrimination Training

The basic procedure for **stimulus discrimination training** entails a multiple schedule (Chapter 13) with antecedent stimulus conditions representing each component schedule. Responses in the presence of one stimulus condition (S^D) are reinforced, and responses in the presence of the other stimulus (S^Δ) are not reinforced. When this procedure is applied appropriately and consistently, responding in the presence of the S^D will come to exceed responding in the presence of the S^Δ. Often, over time, the participant will learn not to respond in the presence of the S^Δ.

Applied behavior analysts often describe the conventional procedures for discrimination training, with differential reinforcement as alternating conditions of reinforcement and extinction, meaning that a response produces reinforcement in the S^D condition but not in the S^Δ condition. However, to clarify and stress and important point: the S^Δ is used not only to show a condition of zero reinforcement (extinction) but also to denote a condition that provides a lesser amount or quality of reinforcement than the S^D condition (Michael, 2004).

Maglieri, DeLeon, Rodriguez-Catter, and Sevin (2000) used discrimination training as part of an intervention to decrease the food stealing of a 14-year-old girl with Prader-Willi syndrome, a serious medical condition usually correlated with obesity and food stealing. During discrimination training, a teacher showed the girl two containers of cookies. One container contained a warning label, the S^Δ, and the other container had no warning label, the S^D. The teacher told the girl that she could eat cookies only from the container without the warning label. The teacher asked the girl in the presence of the two containers, "Which cookies can you eat?" If the girl answered that she could eat the cookies from the container without the warning, the teacher let her eat one cookie from the container. This discrimination training procedure decreased stealing food from the containers marked with the warning label.

Researchers have used multiple schedules to help learners develop stimulus control of social approach responses (e.g., Cammilleri, Tiger, & Hanley, 2008; Grow, LeBlanc, & Carr, 2010; Torelli, Lloyd, Diekman, & Wehby, 2016). Although teaching students to recruit teacher attention can increase teacher praise and instructional assistance and improve student academic performance (see Teach the Learner to Recruit Reinforcement in Chapter 30), classrooms are busy places and teacher attention is not continuously available. Students must be taught not only *how* to recruit teacher attention but also how to discriminate *when* recruiting is appropriate (Alber & Heward, 2000).

Cammilleri et al. (2008) used a classwide multiple-schedule procedure to bring young children's requests to teachers under stimulus control in the three primary classrooms. In two to four 5-min sessions per day throughout the study, the teachers randomly alternated wearing a green lei or a red lei. During baseline, differential consequences were not paired with each lei. The teachers responded to a student's social approach and provided academic assistance as needed. During the multiple-schedule condition, the green lei (S^D) signaled that teacher attention was available, and the red lei (S^Δ) signaled teacher attention was unavailable. Prior to each session during the multiple-schedule condition, the teacher described the contingency associated with the relevant lei: "While I am wearing the green lei, I will be able to answer your questions" or "While I am wearing the red lei, I will not be able to answer your questions" (p. 301). Student requests during the multiple-schedule condition were maintained during desirable periods and minimized during undesirable periods (see Figure 17.3).

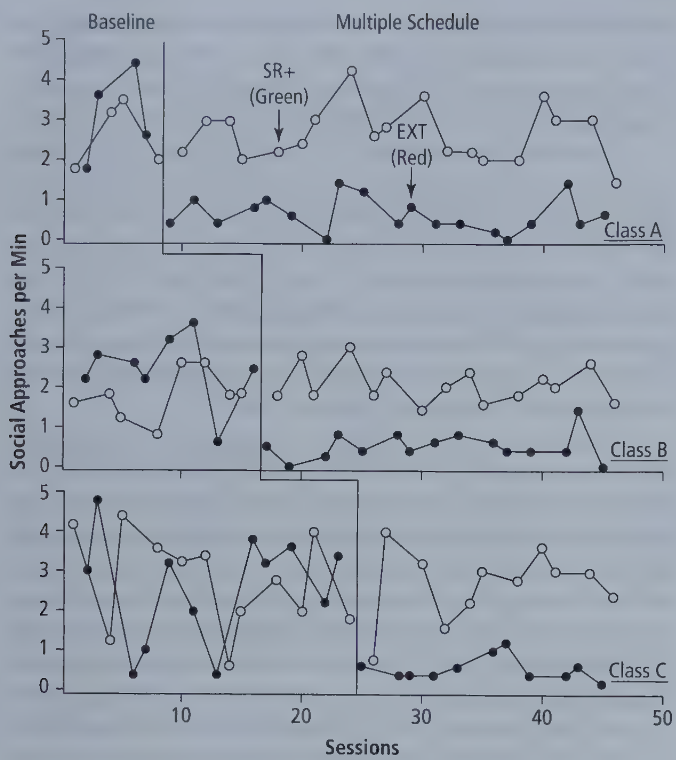


Figure 17.3 Social approach rates during baseline and multiple-schedule conditions across three elementary classrooms.

"Developing Stimulus Control of Young Children's Requests to Teachers: Classwide Applications of Multiple Schedules", A. P. Cammilleri, J. H. Tiger, and G. P. Hanley, 2008, Reproduced with permission of John Wiley & Sons Inc.

Teaching Conditional Discriminations

Skinner conducted early research on complex stimulus control with an experimental procedure called **matching-to-sample**. Dinsmoor (1995b) described Skinner's procedure in the following way. A pigeon was presented with three horizontal keys to peck. The middle key was illuminated with a color to start the trial. A peck on the illuminated key turned it off and lighted two side keys. One of the side keys was the same color as the sample color illuminated on the middle key. A peck on the side key with the same color as the sample key produced reinforcement. Error responses were not reinforced.

Skinner's matching-to-sample procedure included the *three-term contingency*:

S^D ⟶ Response ⟶ Reinforcement
Side key with color | Key peck | Grain

This basic contingency, however, is incomplete because of constraints from the environmental context. The contextual events operating on the three-term contingency become *conditional discriminations* (Sidman, 1994; Spradlin & Simon, 2011). The color of the middle key in Skinner's procedure is the conditional stimulus. The three-term contingency is effective only when it matches the sample stimulus. Reinforcement is conditional on the context of discriminative stimuli other than the S^D; that is, the three-term contingency itself comes under contextual control. Conditional discriminations operate at the level of a *four-term contingency*:

Conditional sample ⟶ S^D ⟶ Response ⟶ Reinforcement
Color of middle key ⟶ Side key same color as middle key ⟶ Key peck ⟶ Grain
S^{Δ} ⟶ Key peck ⟶
Side key different color than middle key

To start a matching-to-sample trial, the participant will make a response (called the observing response) to present the sample stimulus (i.e., the conditional sample). The comparison stimuli (i.e., the discriminative events) are presented usually after removing the sample stimulus, but not always, and provide for an effective three-term contingency and other noneffective three-term contingencies. One comparison stimulus will match the conditional sample. A response that selects the matched comparison and rejects the nonmatched comparisons will produce reinforcement. Figure 17.4 presents an example

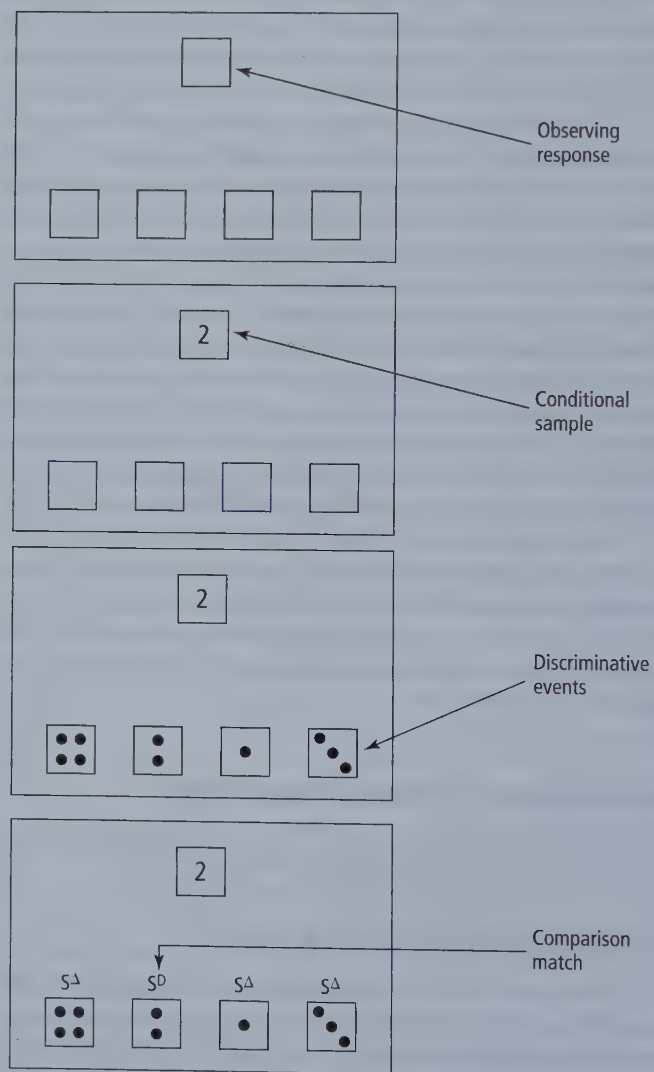


Figure 17.4 An example of the observing response, the conditional sample, the discriminative events, and the comparison match during a matching-to-sample trial.

of the observing response, the conditional sample, the discriminative events, and the comparison match. Responses that select the nonmatching comparison stimuli are not reinforced. During conditional discrimination training, the same selection must be correct with one conditional stimulus, but incorrect with one or more other sample stimuli. Most matching-to-sample applications use a correction procedure with error responses. One correction procedure has the learner respond to the same sample and comparison stimuli until a correct response has been reinforced. Error correction procedures and the random positioning of the comparison stimuli control for position responding. Numerous examples of using the matching-to-sample procedure to teach conditional discriminations are described in Chapter 19.

Teaching Concepts

The preceding section on discrimination training describes how an antecedent stimulus can acquire evocative control over a response, meaning that the behavior occurs more frequently in the presence of the stimulus than in its absence. A discrimination training procedure might be used to teach a preschool student to name the primary colors. For instance, to teach the color red, the teacher could use a red object, such as a red ball, as the S^D condition and a non-red object, such as a yellow ball, as the S^Δ condition. The teacher could position both balls randomly in front of the student, direct the student to name and point to the red ball, and reinforce correct responses, but not incorrect responses. After a few trials the red ball would acquire stimulus control over the student's response, and the student would reliably differentiate the red ball from the yellow ball. This simple discrimination training, however, is unlikely to meet the instructional objective of identifying the color red. The teacher may want the student to learn not only to discriminate between red balls and balls of different color but also the concept of *redness*. A **concept** is defined by a set of shared features found in each example of the concept (Layng, 2017).

> That is, every example of a concept shares certain *must have* features with all other examples of the concept. In addition to these *must have* features, the examples have other *can have* features, which other examples of the concept may or may not have. . . . Thus, the shared *must have* features are what define something as an example of a concept and do not change from example to example. The other nondefining features are those the example *can have*, often changing from example to example, but that don't define the example as an instance of the concept. (Layng, 2017, pp. 1–2)

Terms such as *concept formation* or *concept acquisition* imply for many people some hypothetical construct or mental process. Yet acquiring a concept is clearly dependent on responses in the presence of antecedent stimuli and the consequences following those responses. When given examples and nonexamples of a concept not presented during instruction, a learner who understands the concept can:

- Distinguish examples of the concept from similar nonexamples that lack one or more of the defining *must have* features (discrimination).
- Identify examples of the concept across a wide range of varying examples that include nondefining *can have* features (stimulus generalization). (adapted from Layng, 2017)

Concept formation is a behavioral outcome of stimulus generalization and discrimination (Keller & Schoenfeld, 1950; Zentall, Galizio, & Critchfield, 2002). Concept formation is a complex example of stimulus control that requires both stimulus generalization within a class of stimuli and discrimination between classes of stimuli. An **antecedent stimulus class** is a set of stimuli that share a common relationship. All of the stimuli in an antecedent stimulus class will evoke the same operant response class or elicit the same respondent behavior. This evocative or eliciting function is the only common property among the stimuli in the class (Cuvo, 2000). For example, consider a stimulus class for the concept *red*. A red object is called red because of a particular conditioning history. This conditioning history of differential reinforcement will evoke the response *red* to light waves of different wavelengths, from light red to dark red. These different shades of the color red that will evoke the response *red* share a conditioning history and are included in the same stimulus class. Shades of red (e.g., very light red) that do not evoke the response *red* are not members of that stimulus class. Therefore the concept of redness requires stimulus generalization from the trained stimuli to many other stimuli within the stimulus class. If the preschool student described earlier had acquired the concept of redness, he would be able to identify the red ball and, without specific training or reinforcement, choose a red balloon, a red toy car, a red pencil, and so on.

In addition to stimulus generalization, a concept requires discrimination between members and nonmembers of the stimulus class. For example, the concept of redness requires discriminating between red and other colors and irrelevant stimulus dimensions such as shape or size. The concept begins with discrimination between the red ball and the yellow ball but results in discriminating a red dress from a blue dress, a red toy car from a white toy car, and a red pencil from a black pencil.

Discrimination training is fundamental to teaching conceptual behavior. Antecedent stimuli representative of a group of stimuli sharing a common relationship (i.e., the stimulus class) and antecedent stimuli from other stimulus classes must be presented. Before a concept can be acquired, the teacher must present exemplars of what the concept is (i.e., the S^D condition) and what the concept is not (i.e., the S^Δ condition). This approach holds true for all conceptual development, even for highly abstract concepts (e.g., honesty, patriotism, justice, freedom, sharing). It is also possible to acquire a concept through vicarious discrimination training and differential reinforcement. A verbal definition of a concept, with examples and nonexamples of the concept, may be sufficient for concept formation without additional direct training.

Authors of children's literature often teach concepts vicariously, such as good and bad, honest and dishonest, courageous and cowardly. For example, consider the story of an owner of a mom-and-pop grocery store who wanted to hire a young person to work in the store. The job included sweeping the floor, bagging groceries, and keeping the shelves neat. The owner wanted an honest person to work for her, so she decided to test all applicants to see whether they were honest. The first young person who applied was given the opportunity to try the

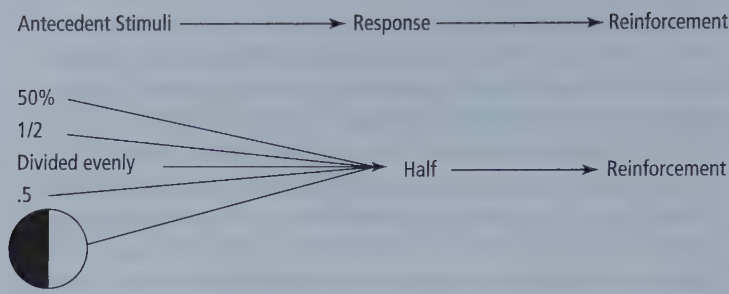


Figure 17.5 Example of an arbitrary stimulus class: Antecedent stimuli with different physical forms that evoke the same response, *half.*

job before the owner made the commitment to hire him. But before the applicant came to work, the owner hid a dollar bill where she knew the young person would find it. At the end of the test period the owner asked the applicant how he had liked working in the store, whether he wanted the job, and whether anything surprising or unusual had happened to him. The applicant replied that he wanted the job and that nothing surprising had happened to him. The grocer told the first applicant that she wanted to consider others who had also applied. The second applicant worked a test period with the same results as the first person. He did not get the job. The third young person to work for the grocer was sweeping the floor, found the dollar bill, and took it immediately to the grocer. The third applicant said he turned in the dollar bill in case one of the customers or the grocer had dropped it. The grocer asked the applicant whether he liked the job and wanted to work for her. The young person replied that he did. The grocer told the applicant that he had the job because he was an honest person. The grocer also let him keep the dollar bill.

The previous children's story presents exemplars of honest and dishonest behavior. The honest behavior was rewarded (i.e., the honest person obtained the job), and the dishonest behavior was not rewarded (i.e., the first two applicants did not get the job). This story might vicariously teach a certain concept of honesty.

Stimuli that make up a class may function within *feature stimulus classes* and *arbitrary stimulus classes* (McIlvane, Dube, Green, & Serna, 1993; Catania, 2013; Spradlin & Simon, 2014). Stimuli in a **feature stimulus class** share common physical forms (e.g., topographical structures) or common relative relations (e.g., spatial arrangements). Feature stimulus classes include an infinite number of stimuli and constitute a large portion of our conceptual behavior. For instance, the concept 'dog' is based on a feature stimulus class. The common physical forms of all dogs will be members of that stimulus class. A young child, through differential reinforcement, will learn to differentiate dogs from horses, cats, cows, and so on. Physical forms provide common relations for many feature stimulus classes, such as stimulus classes that evoke the responses book, table, house, tree, cup, cat, rug, onion, and car. A *relational*, or relative, relation exists among stimuli in other feature classes. Examples of these feature stimulus classes based on relative relations are found with concepts such as bigger than, hotter than, higher than, on top of, and to the left of.

Stimuli composing an **arbitrary stimulus class** evoke the same response, but they do not share a common stimulus feature (i.e., they do not resemble each other in physical form, nor do they share a relational relationship). For example, a teacher could form an arbitrary stimulus class using the stimuli *50%, 1/2, divided evenly, .5*, and a shaded circle (see Figure 17.5). Following training, each of these stimuli of different physical forms will evoke the same response, *half. Green bean, asparagus, potato*, and *corn* could be developed into an arbitrary stimulus class to evoke the response *vegetable*. English language students learn to associate vowels with the arbitrary class of the letters *A, E, I, O, U*, and sometimes *Y*.

The development of concepts and complex verbal relations plays an important role in parenting, caregiving, education, and treatment. Applied behavior analysts need to consider different instructional procedures when teaching concepts and complex verbal relations that produce feature stimulus classes and arbitrary stimulus classes. A common instructional procedure used with feature stimulus classes is to differentially reinforce responses to *is* (S^D) and *is not* (S^Δ) examples of the concept. Broad generalization is common with feature stimulus classes; depending on the functioning level of the participants, a few trained examples may be sufficient to develop the concept. Stimulus generalization, however, is not a characteristic of arbitrary stimulus classes. Applied behavior analysts have developed arbitrary stimulus classes by using matching-to-sample procedures to create stimulus equivalence among arbitrary stimuli (see Chapter 19).

Using Response and Stimulus Prompts

Prompts are supplementary antecedent stimuli used to occasion a correct response in the presence of the natural S^D that will eventually control the behavior. Applied behavior analysts give response and stimulus prompts before or during the performance of a behavior. **Response prompts** operate directly on the response to cue a correct response. The three major forms of response prompts are verbal instructions, modeling, and physical guidance. **Stimulus prompts** operate directly on the antecedent task stimuli to cue a correct response in conjunction with the critical S^D.

Verbal Instructions

Applied behavior analysts use functionally appropriate verbal instructions as supplementary response prompts. Verbal response prompts occur frequently in almost all training contexts in the forms of vocal verbal instruction (e.g., oral, telling) and nonvocal verbal instruction (e.g., written words, manual signs, pictures).

Teachers often use vocal verbal instructional prompts. Suppose a teacher asks a student to read the sentence "Plants

need soil, air, and water to grow." The student reads, "Plants need . . . Plants need . . . Plants need . . . " The teacher could use any number of verbal prompts to occasion the next word. She might say, "The next word is *soil*. Point to *soil* and say *soil*." Or she might use a rhyming word for *soil*. For another example, Adkins and Mathews (1997) taught in-home caregivers to use vocal verbal response prompts to improve voiding procedures for two adults with urinary incontinence and cognitive impairments. The in-home caregiver checked for dryness each hour or every 1 or 2 hours between 6:00 AM and 9:00 PM. The caregiver praised dryness, asked the adult to use the toilet, and provided assistance as needed when the adult was dry at the regularly scheduled check. This simple response prompt procedure (i.e., asked the adult to use the toilet), which was introduced following a baseline condition, produced in one of the adults a mean 22% reduction of grams of urine collected per day in wet diapers during the 2-hour prompted voiding condition, and a mean 69% reduction during the 1-hour prompt condition. The second adult received only the 1-hour prompted voiding condition, which resulted in a mean 55% reduction of urine collected per day.

Krantz and McClannahan (1998) and Sarokoff, Taylor, and Poulson (2001) used nonvocal verbal instructional response prompts in the form of embedded scripts to improve the spontaneous social exchanges of children with autism. Examples of the embedded scripts in the children's photographic activity schedules included *look, watch*, and *let's eat our snack*. In another example of nonvocal verbal instructional response prompts, Wong, Seroka, and Ogisi (2000) developed a checklist with 54 steps to prompt self-assessment of blood glucose level by a diabetic woman with memory impairments. This participant followed the checklist sequence and checked off each step as she completed it.

Modeling

Applied behavior analysts can demonstrate or model the desired behavior as a response prompt. Modeling can effectively prompt behaviors especially for learners who have already learned some of the component behaviors required for the imitation. Modeling is an easy, practical, and successful way for a coach to show a player an appropriate form for shooting a basketball through a hoop when the player already can hold the ball, raise it over his head, and push the ball away from his body. Few teachers would use modeling to teach a child with severe disabilities to tie her shoes if she could not hold the laces in her hands. In addition, attending skills are important. The learner must observe the model to enable imitation of the performance. Finally, modeling as a response prompt should be used only with learners who have already developed imitative skills. The use of models to develop a wide range of new behaviors has been demonstrated repeatedly. Chapter 21 provides a detailed discussion of modeling, imitation, and observational learning.

Physical Guidance

Physical guidance is a response prompt applied most often with young children, learners with severe disabilities, and older adults experiencing physical limitations. Using physical guidance, the teacher partially guides the student's movements or physically guides the student throughout the entire movement of the response.

Hanley, Iwata, Thompson, and Lindberg (2000) reported the use of physical guidance to help participants with profound intellectual disabilities manipulate leisure items. Conaghan, Singh, Moe, Landrum, and Ellis (1992) used physical guidance to prompt adults with intellectual disabilities and hearing impairments to use manual signs. When a participant made an error in sign production, the teacher physically guided the person's hands to prompt a correct response. In another example, a personal trainer worked with three older adults with severe disabilities, osteoporosis, and arthritis. The trainer physically guided the participants' arm movements whenever they did not begin independent pushes with a dumbbell or stopped the pushes before reaching their exercise criterion (K. Cooper & Browder, 1997).

Physical guidance can be an effective response prompt, but it is more intrusive than verbal instruction and modeling. It requires direct physical contact between the teacher and the student, making precise assessment of student progress difficult. Even though some learners require physical guidance, such response prompts provide little opportunity for the student to emit the behavior without the direct assistance of the teacher. Another possible problem is that some learners resist physical touch.

Stimulus Prompts

Applied behavior analysts frequently use movement, position, and redundancy of antecedent stimuli as stimulus prompts. For example, movement prompts can help a learner discriminate between a penny and a dime by pointing to, tapping, touching, or looking at the coin to be identified. In the coin discrimination task, the teacher could use a position prompt and place the correct coin closer to the student. Redundancy prompts occur when one or more stimulus or response dimensions (e.g., color, size, shape) are paired with the correct choice. For instance, a teacher might use a color mediation procedure of associating a numeral with a color, then link the name of a color to an answer for an arithmetic fact (Van Houten & Rolider, 1990).

Practitioners must be aware that there is a risk of prompt dependence with the applications of response and stimulus prompts (e.g., Grow & LeBlanc, 2013). Prompt dependence results from a failure to transfer stimulus control from the prompt to the desired stimuli. The basic research foundations and procedures for the transfer of stimulus control are presented later in this chapter.

Factors Affecting the Development of Stimulus Control

Applied behavior analysts establish stimulus control with frequent differential reinforcement of behavior in the presence and absence of the S^D condition. Effective differential reinforcement requires the consistent use of consequences that function as reinforcers. Additional factors such as preattending skills, stimulus salience, masking, and overshadowing also will affect the development of stimulus control.

Preattending Skills

The development of stimulus control requires certain prerequisite skills. For academic or social skills, the student should engage in orienting behaviors appropriate to the S^Ds in the instructional setting. Such preattending skills include looking at the instructional materials, looking at the teacher whenever a response is modeled, listening to oral instructions, and sitting quietly for short periods. Teachers should use direct behavioral interventions to specifically teach preattending skills to learners who have not developed these skills. Learners must emit behaviors that orient their sensory receptors to the appropriate S^D for the development of stimulus control.

Stimulus Salience

The salience of a stimulus (i.e., the prominence of the stimulus in the learner's environment) influences attention to it and ultimately the development of stimulus control (Dinsmoor, 1995b). For example, Conners and colleagues (2000) included salient cues (e.g., specific room color, specific therapist) into the context of multielement functional analyses (see Chapter 27). Their results suggested that salient cues facilitated the efficiency of functional analyses by producing faster and clearer outcomes in the presence of the salient cues than in the absence of them.

Some stimuli have more salience than others, depending on the sensory capabilities of the individual, the past history of reinforcement, and the context of the environment. For instance, a student may not attend to words written on the blackboard because of poor vision, or to the teacher's oral directions because of poor hearing, or to curriculum materials because of past failures to learn, or to teacher directions because of focusing attention on a toy in the student's desk.

Catania (2013) clarified the common use of the term *salience* (as we have used it here) when he stated that "salience isn't a property of a stimulus; it is actually a property of the organism's behavior with respect to that stimulus" (p. 140). We can say this another way: Although a behavior analyst can change various aspects (e.g., size, position, color) of a stimulus in an effort to increase its salience, whether or not salience is achieved is determined by the learner's differential responding (or lack thereof) to those changes.

Overselective Stimulus Control

In **overselective stimulus control** (also called *stimulus overselectivity*) the range of discriminative stimuli or stimulus features controlling behavior is extremely limited (Lovaas, Koegel, & Schreibman, 1979). Some children with autism or intellectual disabilities focus on a minute feature of an object or a person rather than the whole (Dube et al., 2016; Ploog, 2010). For example, if shown a guitar for the first time, a child might focus on the sound hole and not consider anything else about the instrument, such as its size, shape, other parts, or even the sound it makes. This selective stimulus control interferes with the child's understanding of what a guitar is—the totality of its parts and function. Overselective stimulus control hinders a child's ability to interpret relevant meaning from the environment and to learn new concepts and academic and language skills.

Several studies have demonstrated that teaching learners to emit more elaborate observing responses before making a selection response can improve the performance of learners in matching-to-sample activities (Doughty & Hopkins, 2011; Dube & McIlvane, 1999; Gutowski & Stromer, 2003). The elaborated observing response used by Farber, Dickson, and Dube (2017) to improve the performance of participants with overselective stimulus control was described as follows:

> In typical matching procedures, the sample stimuli are presented first, the student makes an observing response to the samples, and then an array of comparison stimuli is presented, one of which is correct in relation to the samples. Observing responses to sample stimuli may be classified as nondifferential or differential. With nondifferential observing responses, the response to samples is the same on every trial and the only discrimination required is between the presence versus absence of the stimuli. By contrast, differential observing responses (DORs) include behavioral requirements that verify discrimination of stimulus features that differ among the samples. For example, if the sample stimuli are pictures, one type of DOR would be to name the pictures aloud (e.g., Constantine & Sidman, 1975). (Farber et al., 2017, p. 88)

Stimulus Blocking and Overshadowing

Stimulus blocking and overshadowing are methods for decreasing the salience of stimuli (Dinsmoor, 1995b). In **stimulus blocking** (sometimes called *masking*), even though one stimulus has acquired stimulus control over behavior, a competing stimulus can block the evocative function of that stimulus (Didden, Prinsen & Sigafoos, 2000; Seraganian & vom Saal, 1969). For example, a student may know the answers to a teacher's questions, but will not respond in the presence of the peer group, which in this example is competition of different contingencies of reinforcement, not just an antecedent stimulus that makes it harder to "attend to" the relevant S^D. In **overshadowing**, the most salient component of a compound stimulus arrangement controls responding and interferes with the acquisition of stimulus control by the more relevant stimulus. For example, Dittlinger and Lerman (2011) found that pictures in text delayed acquisition of sight word reading by children with developmental disabilities, suggesting that pictures overshadowed the text and may have interfered with learning. Fading in the pictures using a least-to-most prompting hierarchy and a text-to-picture matching response requiring students to attend to both the text and the picture have proved effective against overshadowing (Richardson et al., 2017).

Applied behavior analysts need to recognize that blocking and overshadowing can hinder the development of stimulus control and apply procedures that reduce these effects. Examples of reducing the influence of blocking and overshadowing include (a) rearranging the physical environment (e.g., lowering the window shade, removing distractions, changing seating assignments), (b) making instructional stimuli appropriately intense (e.g., providing a rapid pace of instruction, many opportunities to respond, an appropriate level of difficulty, and opportunities

to set goals), and (c) consistently reinforcing behavior in the presence of the instructionally relevant discriminative stimuli.

TRANSFERRING STIMULUS CONTROL

Applied behavior analysts should provide response and stimulus prompts as *supplementary* antecedent stimuli only during the acquisition phase of instruction. With the reliable occurrence of behavior, applied behavior analysts need to transfer stimulus control from the response and stimulus prompts to the naturally existing stimulus. Applied behavior analysts transfer stimulus control by gradually fading stimuli in or out, gradually presenting or removing antecedent stimuli. Eventually, the natural stimulus, a partially changed stimulus, or a new stimulus will evoke the response. Fading response prompts and stimulus prompts is used to transfer stimulus control from the prompts to the natural stimulus, and also to minimize the number of error responses occurring in the presence of the natural stimulus. For a review of response-prompt and stimulus-prompt fading procedures, see Cenghar, Budd, Farrell, and Fienup (2018).

Basic Research Foundations

Terrace's (1963a, b) influential research on the transfer of stimulus control using fading and superimposition of stimuli provides classic examples of transferring stimulus control. In these studies, Terrace taught pigeons to make red–green and vertical–horizontal discriminations with a minimum of errors. His use of techniques for gradually transferring stimulus control was called **errorless learning**. To teach a red–green discrimination, Terrace presented the S^{Δ} (red light) at the beginning of discrimination training, before the S^D (green light) had stimulus control over the pigeon's responses. The initial introduction of the red light was with low illumination and for brief time intervals. During successive presentations of the stimuli, Terrace gradually increased the intensity of the red light and the duration it was illuminated until it differed from the green light only in hue. With this procedure, Terrace taught pigeons to discriminate red from green with only a minimum number of errors (responses to the S^{Δ}).

Terrace further demonstrated that stimulus control acquired with red and green lights could be transferred to vertical and horizontal lines with a minimum number of errors (i.e., responses in the presence of the S^{Δ}). His procedure consisted of first superimposing a white vertical line on the green light (S^D) and a white horizontal line on the red light (S^{Δ}). Then the pigeons were given several presentations of the two compound stimuli. Finally, the amplitude of the red and green lights was reduced gradually until only the vertical and horizontal lines remained as stimulus conditions. The pigeons showed almost perfect transfer of stimulus control from the red–green lights to the vertical–horizontal lines. That is, they emitted responses in the presence of the vertical line (S^D) and seldom responded in the presence of the horizontal line (S^{Δ}).

Following Terrace's work, other pioneer researchers (e.g., Moore & Goldiamond, 1964) produced landmark studies showing that the transfer of stimulus control with few incorrect responses was possible with human learners, which provided the foundations for developing effective procedures to transfer stimulus control from response prompts to natural stimuli in the applied context.

Transferring Stimulus Control from Contrived Response Prompts to Naturally Existing Stimuli

Wolery and Gast (1984) described four procedures for transferring stimulus control from response prompts to natural stimuli. They describe these procedures as most-to-least prompts, graduated guidance, least-to-most prompts, and time delay.

Most-to-Least Prompts

The applied behavior analyst can use most-to-least response prompts to transfer stimulus control from response prompts to the natural stimulus whenever the participant does not respond to the natural stimulus or makes an incorrect response. To apply **most-to-least response prompts**, the analyst physically guides the participant through the entire performance sequence, and then gradually reduces the amount of physical assistance provided as training progresses from trial to trial and session to session. Customarily, most-to-least prompting moves from physical guidance to visual prompts to verbal instructions, and finally to the natural stimulus without prompts.

Graduated Guidance

The analyst provides physical guidance as needed, but using graduated guidance she will immediately start to fade out the physical prompts to transfer stimulus control. Graduated guidance begins with the applied behavior analyst following the participant's movements closely with her hands, but not touching the participant. The analyst then increases the distance of her hands from the participant by gradually changing the location of the physical prompt. For example, if the analyst used physical guidance for a participant's hand movement in zippering a coat, she might move the prompt from the hand to the wrist, to the elbow, to the shoulder, and then to no physical contact. Graduated guidance provides the opportunity for an immediate physical prompt as needed.

Least-to-Most Prompts

When transferring stimulus control using **least-to-most response prompts**, the applied behavior analyst gives the participant an opportunity to perform the response with the least amount of assistance on each trial. The participant receives greater degrees of assistance with each successive trial without a correct response. The procedure for least-to-most prompting requires the participant to make a correct response within a set time limit (e.g., 3 seconds) from the presentation of the natural S^D. If the response does not occur within the specified time, the applied behavior analyst presents the natural S^D and a response prompt of least assistance, such as a verbal response prompt. If after the same specified time limit (e.g., another 3 seconds) the participant does not make a correct response, the analyst gives the natural S^D and another response prompt, such as a gesture. The participant receives partial or full physical guidance if the lesser prompting does not evoke a correct response.

Applied behavior analysts using the least-to-most response prompt procedure present the natural S^D with the same time limit for an unprompted correct response each trial. For example, Heckaman, Alber, Hooper, and Heward (1998) used instructions, nonspecific verbal prompts, modeling, and physical prompting in a least-to-most 5-sec response prompt hierarchy with four students with autism.

Based on the results of two studies in which they compared the effectiveness of most-to-least and least-to-most prompting on the acquisition of solitary play skills by children with autism, and previous research, Libby, Weiss, Bancroft, and Ahearn (2008) recommended the following guidelines for selecting prompting techniques:

- Most-to-least should be the default prompting technique when a child's learning history is unknown.
- If errors have been found to evoke problem behavior or impede learning, most-to-least or most-to-least with time delay is preferable. (Time delay is described in the following section.)
- Least-to-most may be preferable for students who have shown rapid skills acquisition with this technique.
- In all cases, progress should be monitored to ensure errors do not impede learning. (adapted from Libby et al., 2008)

Time Delay

Touchette (1971) discovered that inserting small increments of time between presentation of the instructional stimulus and a response prompt across trials resulted in the learner "anticipating" the correct response. A **time delay** procedure begins with simultaneous presentation of the target stimulus (e.g., picture of a dog) and response prompt (e.g., teacher says, "dog"). After the student has responded correctly for several trials, the teacher inserts a delay between the instructional stimulus and the response prompt until the student emits the unprompted correct response. Time delay is considered an "errorless learning" technique because students often make few or no errors as control of responding shifts from the contrived prompt to the instructional stimulus.

Two variations of time delay are progressive time delay and constant time delay. A **progressive time delay** procedure starts with simultaneous presentation of the natural stimulus and the response prompt (i.e., 0-sec delay). Usually, a teacher will use several 0-sec trials before extending the time delay. The number of 0-sec trials will depend on the task difficulty and the functioning level of the participant. Following the simultaneous presentations, the teacher gradually and systematically extends the time delay, often in 1-sec intervals. With **constant time delay**, after the student has responded correctly to several 0-sec delay trials, presentation of the response prompt follows the instructional stimulus by a predetermined and fixed delay (usually 3 or 4 seconds) for all subsequent trials.

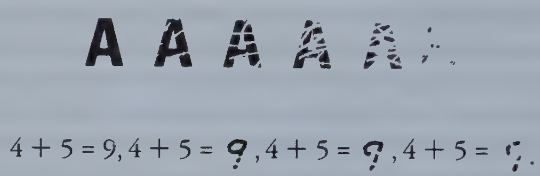
Time delay has been used to teach a variety of skills, such as sight word reading (Knight et al., 2003); spelling (Stevens & Schuster, 1987); numeral identification (Ault et al., 1988); and chained tasks such as food preparation, purchasing, vocational, and leisure skills (Dogoe & Banda, 2009). Time delay has been used effectively by peer tutors (Hughes & Fredrick, 2006) and parents (Dipipi-Hoy & Jitendra, 2004). Heckaman et al. (1998) reported that implementing progressive time delay immediately reduced rates of disruptive behavior associated with a task previously taught with least-to-most prompting procedure. For reviews of research on time delay, see Browder, Ahlgrim-Delzell, Spooner, Mims, and Baker (2009) and Dogoe and Banda (2009).

Transferring of Stimulus Control with Fading Techniques

The preceding sections focused on response prompts that do not change the task stimuli or materials. The stimulus control shaping procedures presented here modify the task stimuli or materials gradually and systematically to prompt correct responses. Supplementary stimulus conditions are faded in or faded out to transfer stimulus control from the stimulus prompt to the natural stimulus. Stimulus control shaping can be accomplished with stimulus fading and stimulus shape transformations (McIlvane & Dube, 1992; Sidman & Stoddard, 1967).

Stimulus Fading

Stimulus fading involves highlighting a physical dimension (e.g., color, size, position) of a stimulus to increase the likelihood of a correct response. The highlighted or exaggerated dimension is faded gradually in or out. The following examples of (a) writing the uppercase letter A and (b) writing the answer 9 to an arithmetic problem are illustrations of stimulus fading.



Krantz and McClannahan (1998) faded out scripts (i.e., the words *Look* and *Watch me*) embedded in photographic activity schedules. The embedded scripts prompted the social exchanges of children with autism. The words *Look* and *Watch me* were printed with 72-point font and bold letters on white 9-cm note cards. Krantz and McClannahan began fading out the words by removing one third of the script card, then another third. Sometimes, during the script fading, portions of letters were still shown on the card, such as part of an *o* in *Look*. Finally, the scripts and cards were removed.

The treatment of a feeding disorder reported by Patel, Piazza, Kelly, Ochsner, and Santana (2001) provides an example of fading stimuli in and out. Severe food selectivity is a common problem for children with feeding disorders. For example, some children find highly textured foods aversive because they evoke a gagging response. Patel and colleagues faded in Carnation Instant Breakfast (CIB) and then milk to water in treating a 6-year-old boy with pervasive developmental feeding disorder. The boy would drink small amounts of water. The researchers began the fading in procedure by adding 20% of the CIB packet to 240 ml of water. Following three sessions of drinking the 20% mixture, more CIB was gradually added to the water, initially by 5% and later 10% increments. The researchers then faded in milk to the CIB and water mixture after the boy would drink one CIB packet with 240 ml of water. Milk was gradually added to

the CIB/water mixture in 10% increments as the water was faded out (e.g., 10% milk and 90% water, plus one packet of CIB; then 20% milk and 80% water).

Applied behavior analysts have used the superimposition of stimuli with stimulus fading. In one instance, the transfer of stimulus control occurs when one stimulus is faded out; in another application, one stimulus is faded in as the other stimulus is faded out. The research by Terrace (1963a, b) demonstrating the transfer of stimulus control from a red–green discrimination to a vertical–horizontal discrimination shows the superimposition of two specific classes of stimuli and the fading out of one stimulus class. The lines were superimposed on the colored lights; then the lights were gradually faded out, leaving only the vertical and horizontal lines as the discriminative stimuli. Figure 17.6 provides an applied example of the superimposition and stimulus fading procedures used by Terrace. The figure shows a series of steps from an arithmetic program to teach 7 − 2 = _____.

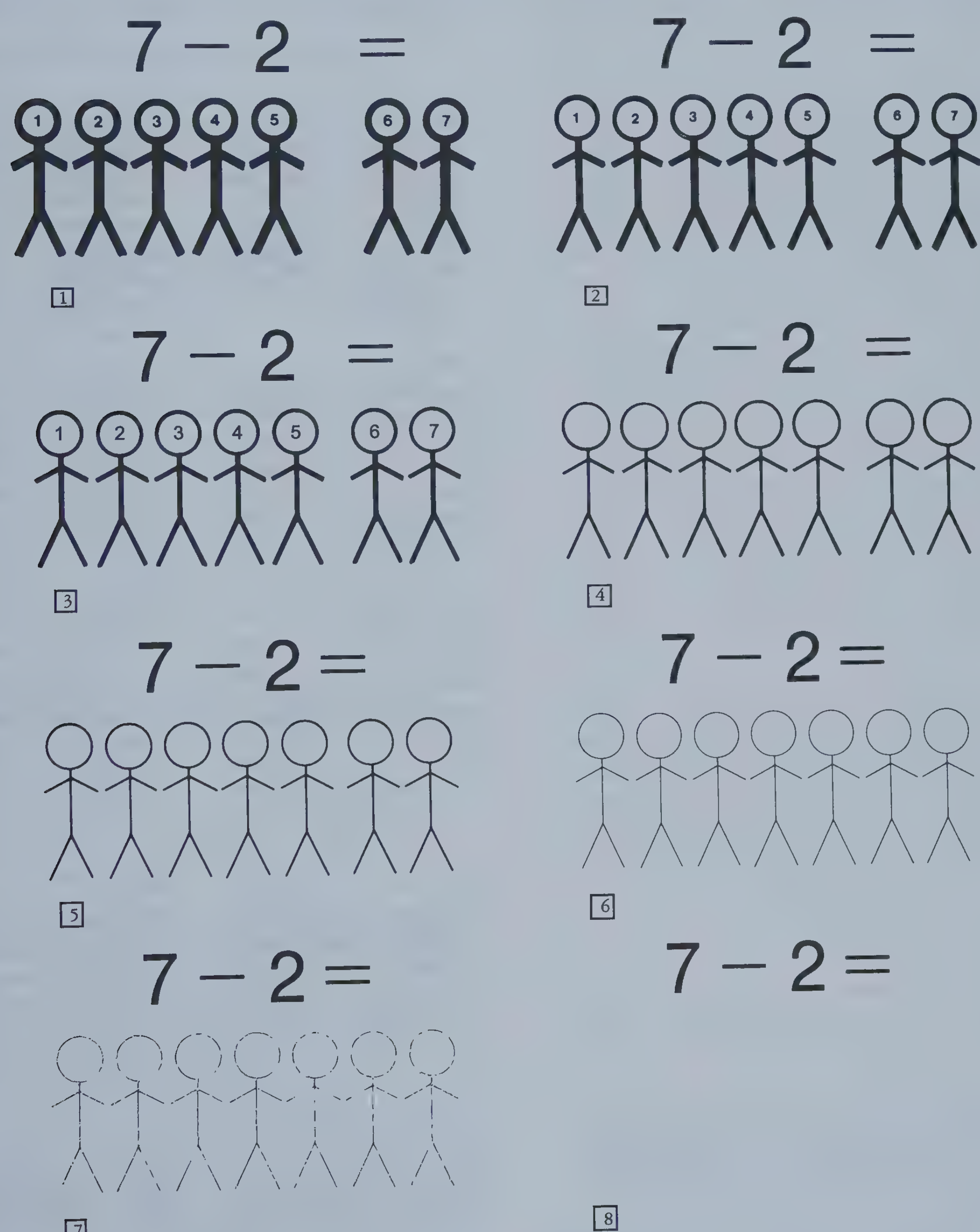
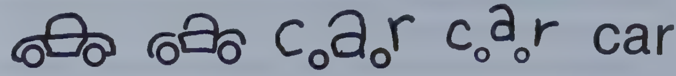


Figure 17.6 Illustration of two classes of superimposed stimuli with one class then faded out.
From *Addition and Subtraction Math Program with Stimulus Shaping and Stimulus Fading*, by T. Johnson, 1973, unpublished project, Ohio Department of Education. Reprinted by permission.

The other frequently used procedure fades in the natural stimulus and fades out the stimulus prompt. Figure 17.7 illustrates this superimposition procedure, in which the prompt is faded out and the natural stimulus 8 + 5 = _____ is faded in.

Stimulus Shape Transformations

The procedure for stimulus shape transformations uses an initial stimulus shape that will prompt a correct response. That initial shape is then gradually changed to form the natural stimulus, while maintaining correct responding. The shape of the stimulus prompt must change gradually so that the student continues to respond correctly. In teaching word identification, using stimulus shape transformations could include the following steps (Johnson, 1973):



In summary, a variety of procedures exist for transferring stimulus control from response and stimulus prompts to natural stimuli. Although fading, superimposition, and stimulus shaping procedures are often successful in transferring stimulus control, preparing the necessary stimulus materials for some tasks requires considerable time and technical skill, and may be costly. Practitioners should always consider the most effective and

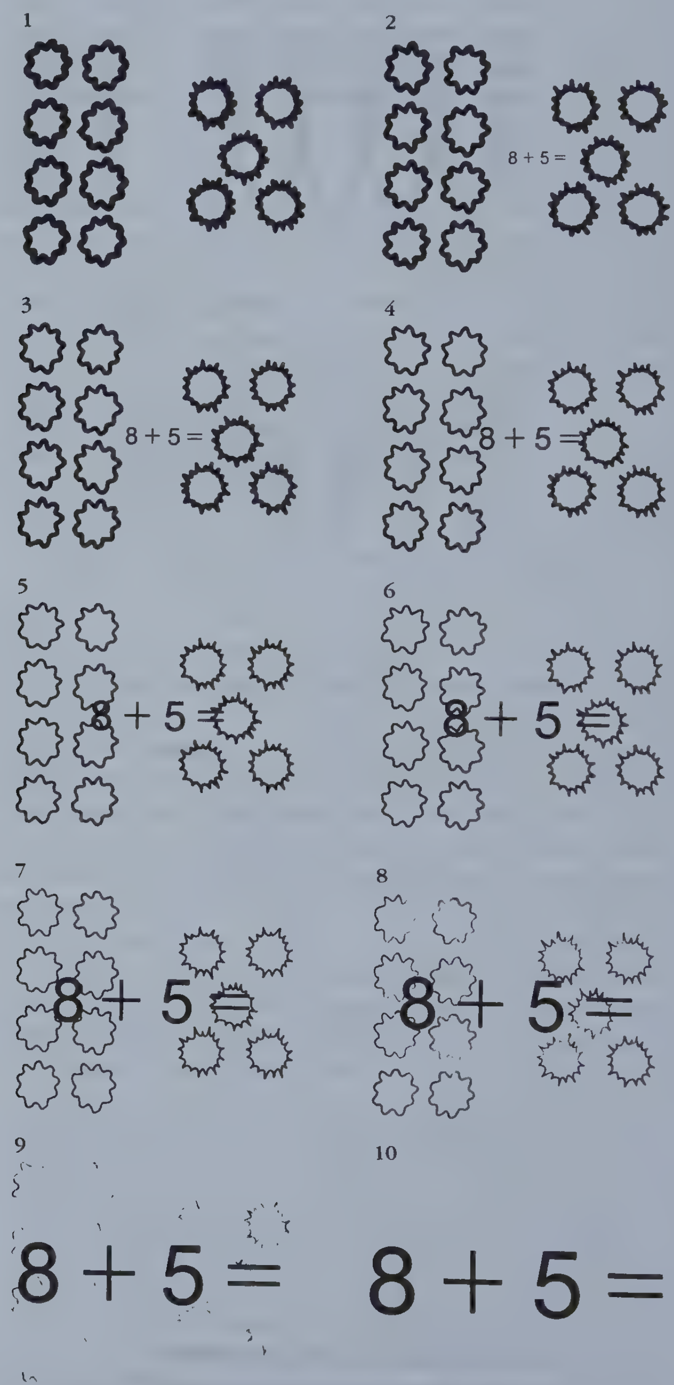


Figure 17.7 Illustration of superimposition and stimulus fading to fade in the natural stimulus and fade out the stimulus prompt.

From Addition and Subtraction Math Program with Stimulus Shaping and Stimulus Fading, by T. Johnson, 1973, unpublished project, Ohio Department of Education. Reprinted by permission.

efficient approach to changing behavior. A study related to binge drinking by college students is informative in this regard. Metz, Kohn, Schultz, and Bettencourt (2017) compared the effectiveness of three training procedures (stimulus fading, superimposition, and verbal feedback) on college students' pouring standard, 12-ounce servings of beer. Four of five participants in the stimulus fading group reached criterion (poured within 10% deviation of 12 oz) without additional training; three of four participants in the superimposition group required and received additional stimulus fading training. But three of four participants in the verbal-feedback-only group achieved criterion on their first set of posttraining pours. The authors concluded that "verbal feedback may be a relatively easy and effective method if used with actual pouring" (p. 29). We believe practitioners should keep in mind this adage: Just because you can do something doesn't mean you should, and always consider the simplest, most efficient methods of behavior change before designing and implementing a complex intervention.

SUMMARY

Stimulus Control: Basic Concepts and Processes

1. Behavior that occurs more often in the presence of a stimulus (discriminative stimulus, S^D) than in its absence (stimulus delta, S^Δ) is said to be under stimulus control.
2. When an S^D signals a rich schedule of reinforcement and a corresponding S^Δ signals a zero chance of reinforcement (i.e., extinction), the organism will eventually respond only in the presence of the S^D. In such cases, stimulus discrimination is complete.
3. Stimulus generalization refers to the extent to which stimuli other than the S^D acquire stimulus control over the behavior. Stimuli sharing similar physical properties with the S^D are most likely to acquire evocative function.
4. A stimulus generalization gradient graphically depicts the degree of stimulus generalization and discrimination by showing the extent to which responses reinforced in one stimulus condition are emitted in the presence of untrained stimuli.
5. Stimulus generalization reflects a loose degree of stimulus control, whereas discrimination indicates a relatively tight degree of stimulus control.
6. In a conditional discrimination, the response that will produce reinforcement in the presence of a particular stimulus depends on (is conditional on) the presence or absence of other stimuli.
7. An S^D acquires its controlling function for operant behavior by being paired with stimulus changes that occur immediately following behavior. Conversely, a conditioned stimulus acquires its controlling function for respondent behavior by being paired with other antecedent stimuli that elicit the behavior (i.e., an unconditioned stimulus or conditioned stimulus).
8. Distinguishing the evocative functions of discriminative stimuli and establishing operations can be difficult. An S^D evokes behavior because its presence has been correlated with the differential availability of an effective reinforcer. EO evokes behavior because it changes what functions as reinforcement rather than signaling the differential availability of reinforcement.

Developing Stimulus Control

9. Stimulus discrimination training entails reinforcing responses in the presence of the S^D and not reinforcing responses in the presence of the S^Δ.
10. Behavior analysts often teach and study conditional discriminations with a matching-to-sample procedure entailing a four-term contingency of observing response, the conditional sample, the discriminative events, and the comparison match.
11. A concept is defined by a set of shared features found in each example of the concept.
12. Concept formation is a complex example of stimulus control that requires both stimulus generalization within a class of stimuli and discrimination between stimulus classes.
13. An antecedent stimulus class is a set of stimuli that share a common relationship. All of the stimuli in the class will evoke the same operant response class, or elicit the same response in the case of respondent behavior.
14. Stimuli in a feature stimulus class share common physical forms (e.g., topographical structures) or common relative relations (e.g., spatial arrangements).
15. Stimuli composing an arbitrary stimulus class evoke the same response, but they do not share a common stimulus feature (i.e., they do not resemble each other in physical form, nor do they share a relational relationship).
16. Prompts are supplementary antecedent stimuli used to occasion a correct response in the presence of an S^D that will eventually control the behavior.
17. Response prompts (e.g., verbal instructions, modeling, physical guidance) operate directly on the response. Stimulus prompts (e.g., movement, position, redundancy) operate directly on antecedent task stimuli to cue a correct response in conjunction with the critical S^D.
18. Applied behavior analysts establish stimulus control with frequent differential reinforcement of behavior in the presence and absence of the S^D condition. Effective differential reinforcement requires the consistent use of consequences that function as reinforcers. Additional factors such as preattending skills, stimulus salience, overselective stimulus control, stimulus blocking, and overshadowing will also affect the development of stimulus control.

Transferring Stimulus Control

19. Terrace (1963a, b) conducted influential basic research using fading and superimposition of stimuli to transfer stimulus control.

20. Fading response prompts and stimulus prompts can transfer stimulus control from the prompts to the natural stimulus and minimize errors by the learner.

21. In the most-to-least response prompting procedure, the analyst physically guides the participant through the entire performance sequence and then gradually reduces the amount of physical assistance provided as training progresses from trial to trial and session to session.

22. Graduated guidance begins with the applied behavior analyst following the participant's movements closely with her hands, but not touching the participant. The analyst then increases the distance of her hands from the participant by gradually changing the location of the physical prompt.

23. With least-to-most prompting, the learner is given the opportunity to perform the response with the least amount of assistance on each trial. The teacher provides increasingly greater degrees of assistance with each successive trial, as needed.

24. Time delay procedures begin with simultaneous presentation of the target stimulus and response prompt. After the student has responded correctly for several trials, the teacher inserts a delay between the instructional stimulus and the response prompt.

25. Stimulus fading involves exaggerating a dimension of a stimulus to increase the likelihood of a correct response. The exaggerated dimension is faded gradually in or out.

26. Stimulus shape transformations begin with an initial stimulus shape that will prompt a correct response. The initial shape is then gradually changed to form the natural stimulus, while maintaining correct responding.

KEY TERMS

antecedent stimulus class
arbitrary stimulus class
concept
concept discrimination
constant time delay
discriminative stimulus (S^D)
errorless learning
feature stimulus class
least-to-most response prompts
most-to-least response prompts
overselective stimulus control
overshadowing
progressive time delay
response prompts
stimulus blocking
stimulus control
stimulus delta (S^Δ)
stimulus discrimination
stimulus discrimination training
stimulus fading
stimulus generalization
stimulus generalization gradient
stimulus prompts
time delay

MULTIPLE-CHOICE QUESTIONS

1. When a group of stimuli all evoke the same response (for example, when one sees a picture of a Border collie, a Doberman pinscher, and a Labrador retriever, one says "that's a dog"), the group of stimuli is referred to as:

a. A response class
b. A stimulus class
c. An arbitrary class
d. A response class

Hint: (See "Developing Stimulus Control")

2. Which of the following is an example of an arbitrary stimulus class?

a. The following women's restroom signs: the word woman, a pictogram of a woman, and a painting of a mermaid
b. A picture of a brick house, a stick built house, and a log cabin
c. A picture of a chair with no arms, a chair with arms, and a rocking chair
d. A picture of a bus, a car, and a motorcycle

Hint: (See "Developing Stimulus Control")

3. Which of the following illustrates an example of stimulus control?

a. Joanne occasionally runs the stop sign near her house because there are rarely any cars at the intersection
b. Frank flaps his hands almost constantly all day long.
c. Joe hits his teacher every time and only when she asks him to wash his hands.
d. Molly uses a Kleenex to wipe her nose when it is running and often when it is not running. It is a bad habit she has.

Hint: (See "Introduction to the chapter")

4. Which of the following statements is true?
   a. When a dog has been trained to salivate when a bell rings by repeatedly pairing meat powder and a bell, this is an example of stimulus control.
   b. When an individual has been trained to say the word "red" when he is shown a card that says RED (but not when he has been shown a card that says GREEN), this is an example of stimulus control.
   c. Antecedent stimuli rarely acquire control over human behavior.
   d. The first two statements are true.

   Hint: (See "Respondent Eliciting and Operant Discriminative Functions of Antecedent Stimuli")

5. A teacher makes a request of a student. The student emits aggressive behavior immediately following the request. The teacher removes the request in order to avoid getting hit. The teacher's request:
   a. Was clearly inappropriate
   b. Is a discriminative stimulus for the problem behavior
   c. Is a motivating operation for the problem behavior
   d. Is a stimulus delta for problem behavior

   Hint: (See "Respondent Eliciting and Operant Discriminative Functions of Antecedent Stimuli")

6. Motivating operations and discriminative stimuli:
   a. Have similar evocative effects on behavior
   b. Are considered synonymous
   c. Occur after the behavior of interest
   d. All of these

   Hint: (See "Discriminative and Motivational Functions of Antecedent Stimuli")

7. Amanda, a typically developing two-year-old girl, has a father who is in the Marines. The last time her father was home, he was in uniform. Now, every time she sees a man in uniform, she says, "Daddy!" This is an example of:
   a. Stimulus control
   b. Stimulus equivalence
   c. A response class
   d. Stimulus generalization

   Hint: (See "Stimulus Generalization")

8. Stimulus generalization is more likely to occur:
   a. When stimuli share more similar properties than when stimuli are very different from one another.
   b. When the practitioner tightly controls the training procedures and stimuli used during training (i.e., using few different stimuli during training).
   c. If time delay is used as part of the training procedure.
   d. If reflexivity is observed.

   Hint: (See "Stimulus Generalization")

9. Stimulus discrimination is acquired by:
   a. Reinforcing responses in the presence of a stimulus delta and withholding reinforcement in the presence of the discriminative stimulus.
   b. Reinforcing responses in the presence of both discriminative stimuli and stimulus deltas.
   c. Reinforcing responses in the presence of the discriminative stimulus and withholding reinforcement in the presence of the stimulus deltas.
   d. Thinning reinforcement so that no reinforcement is required for any antecedent stimuli.

   Hint: (See "Developing Stimulus Control")

10. Stimulus discrimination involves the behavioral principles of:
    a. Prompts and extinction
    b. Reinforcement and extinction
    c. Form and function of behavior
    d. Concepts and stimulus equivalence

    Hint: (See "Developing Stimulus Control")

11. In order to form the concept of "vehicles," an individual must be able to:
    a. Have an IQ of at least 100
    b. Have a strong grasp of concrete operations first
    c. Discriminate vehicles from nonvehicles and generalize across all motorized methods of transporting people
    d. Discriminate vehicles from nonvehicles

    Hint: (See "Developing Stimulus Control")

12. Which of the following is a concept that can be taught using stimulus discrimination and generalization procedures?
    a. Fruit
    b. Under
    c. Integrity
    d. All of these

    Hint: (See "Developing Stimulus Control")

13. The difference between a response prompt and a stimulus prompt is:
    a. A response prompt operates directly on the response, while a stimulus prompt operates on the antecedent task stimuli.
    b. Stimulus prompts are easier to fade than response prompts.
    c. A response prompt involves changing the required response, while a stimulus prompt requires changing the antecedent task stimuli.
    d. Response prompts are verbal, while stimulus prompts are nonverbal.

    Hint: (See "Using Response and Stimulus Prompts")

14. Rahul keeps forgetting to push the Total button on the cash register at his place of employment. His job coach wants to insert a prompt to help him remember to total customers' orders. She wants to use a stimulus prompt. Which of the following is an example of a stimulus prompt?
    a. When Rahul reaches the end of the order, the job coach gives him a nudge on the elbow to get his hand moving toward the Total button.
    b. When Rahul reaches the end of the order, the job coach says, "Don't forget to push Total."
    c. The job coach places a sign on the cash register that says "Don't forget to push Total."
    d. The job coach places a bright pink sticker on the total button.
    Hint: (See "Using Response and Stimulus Prompts")

15. Stimulus and response prompt fading is used to:
    a. Transfer stimulus control from the prompt to the natural antecedent cue.
    b. Thin the reinforcement schedule for responding correctly.
    c. Establish stimulus equivalence.
    d. Promote stimulus generalization.
    Hint: (See "Transferring Stimulus Control")

16. Most-to-least prompt fading involves:
    a. Gradually decreasing the time between the presence of the antecedent cue to the prompt from most time to least time.
    b. Gradually changing the form or intensity of the prompt from most intense to least intense.
    c. Gradually decreasing exaggerated dimensions of the stimulus prompt.
    d. Gradually changing the shape of the prompt to look more like the natural antecedent cue.
    Hint: (See "Transferring Stimulus Control")

ESSAY-TYPE QUESTIONS

1. Explain the basic teaching procedure for helping develop a stimulus class.
   Hint: (See "Developing Stimulus Control")

2. By what mechanisms do antecedents acquire control over respondent and operant behaviors?
   Hint: (See "Respondent Eliciting and Operant Discriminative Functions of Antecedent Stimuli")

3. Brian hits his teacher when he is asked to brush his teeth. However, he never hits his mother when she asks him to brush his teeth. His mother used to have this problem (of Brian hitting her), but she ignored it and made him continue brushing his teeth. His teacher, however, gets frustrated when Brian hits her and usually lets him out of the toothbrushing task (because she thinks it is not her job to get beaten up over doing a task that his parents should have responsibility for anyway). In this scenario, what is the discriminative stimulus and what is the motivating operation for Brian's hitting?
   Hint: (See "Discriminative and Motivational Functions of Antecedent Stimuli")

4. Compare and contrast stimulus generalization and stimulus discrimination.
   Hint: (See "Stimulus Generalization")

5. Explain (or use diagrams to illustrate) how you could teach an individual to say "ball" in the presence of balls, but not in the presence of other stimuli (say bats, blocks, paper).
   Hint: (See "Developing Stimulus Control")

6. Give an example of a concept and delineate what stimuli a learner might need to generalize across and what stimuli the learner might need to discriminate.
   Hint: (See "Developing Stimulus Control")

7. Swapna often forgets to pick up her dirty clothes in the morning and to put them in the clothes hamper. Give an example of each type of response prompt you could use to get her to put her clothes in the hamper. Also, give an example of how you could use stimulus prompts to help her.
   Hint: (See "Using Response and Stimulus Prompts")

8. Give an example of how you could use stimulus fading to help a learner draw a circle.
   Hint: (See "Transferring Stimulus Control")

PART 7

Verbal Behavior

Chapter 18 is devoted to a distinguishing feature of the human behavioral repertoire. Verbal behavior is the principal phenomenon that makes humans interesting and the vehicle with which we express that interest. Verbal behavior makes progress possible from one generation to the next and enhances the development of sciences, technologies, literatures, and the arts. Building on Skinner's (1957) conceptual analysis, Mark Sundberg presents verbal behavior within the context of typical human development, emphasizing the form and function of language, the defining characteristics of verbal behavior, the distinction between topography-based and selection-based verbal behavior, elementary verbal operants and listener discriminations, private events, generative verbal learning, and assessment and intervention programs for children with developmental disabilities.

CHAPTER 18

Verbal Behavior

Mark L. Sundberg

LEARNING OBJECTIVES

- Differentiate between formal properties and functional properties of language.
- Define verbal behavior.
- Define a verbal operant.
- Explain why Skinner suggested that the listener's role is less significant than typically assumed.
- Using a given example, determine the classification of verbal operants and explain how these terms can be used in the analysis of complex verbal behavior.
- Identify and discuss three or more functional units of verbal behavior that contribute to our understanding of multiple control.
- Define an autoclitic relation.
- Identify and discuss how viewing language as a learned behavior involving a social interaction between speakers and listeners, with the verbal operants as the basic units, changes how clinicians and researchers approach and ameliorate human problems related to language.

Jason was a 5-year-old boy diagnosed with autism. He could not talk, echo sounds, or imitate actions. He often screamed to get his needs and wants known. The immediate priorities were to teach Jason to communicate and to reduce his screaming behavior. Jason liked a variety of foods and toys and was taught to ask for them using sign language. He learned his first prompt-free sign ("apple") within 10 minutes of intensive teaching, and his second sign ("Play-Doh") about an hour later. In addition, his screaming behavior was reduced to a near-zero level during that first teaching session. After 4 months of intervention, Jason acquired 43 signs, learned to echo and imitate others, and began to speak. Within 2 years Jason could vocally identify over 300 items and actions, and his problem behaviors were minimal. The intervention program used with Jason was based on applied behavior analysis and Skinner's (1957) analysis of verbal behavior.

This chapter contains two major sections. The first describes the core elements of Skinner's (1957) analysis of verbal behavior, and the second describes some applications of that analysis to language assessment and intervention for persons with autism or other intellectual disabilities.

The author thanks Dave Palmer, Bob Ryan, Hank Schlinger, Carl Sundberg, and Cindy Sundberg for their comments on an earlier version of this chapter.

SKINNER'S (1957) ANALYSIS OF VERBAL BEHAVIOR

B. F. Skinner spent over two decades developing his analysis of language. He began the project a few years after graduating from Harvard University in 1931 and finished it in 1957. His main premise was that language constitutes learned behavior acquired and maintained by environmental contingencies. Reflecting on his work on language, Skinner (1978) wrote "*Verbal behavior* . . . will, I believe, prove to be . . . my most important work" (p. 122). Skinner considered language an important topic because verbal behavior is central to many critical aspects of human behavior (e.g., language acquisition, relationships, intelligence, academics).

It has been said that *Verbal Behavior* (1957) is Skinner's most difficult book to understand. Michael (1984) noted that a reader must first understand the basic concepts and principles of behavior analysis before taking on *Verbal Behavior*. However, a more problematic obstacle noted by Michael was the complexity of the subject matter being analyzed. Linguists have tangled with understanding human communication for centuries. Skinner's work can contribute to their efforts in many ways, beginning with a behavioral distinction between linguistic form and linguistic function.

Form and Function of Language

The formal properties of language involve the topography of a verbal response, and the functional properties involve the causes of that response. For example, a child's cry could be the form of verbal response, and why she cried would be the function.

Skinner's (1957) account of language considers both of these elements but differs from traditional treatments in that he provides a behavioral analysis of the functional properties. He identifies the distinction between form and function early in his book:

> Our first responsibility is simple *description*: what is the topography of this subdivision of human behavior? Once that question has been answered in at least a preliminary fashion we may advance to the stage called *explanation*: what conditions are relevant to the occurrences of the behavior—what are the variables of which it is a function? (p. 10).

Form

To account for both form and function, two separate but interlocking classification systems are necessary. The formal properties of language can be classified by a description of the verbal utterances emitted (e.g., speech, sign language, text, icons). For example, with speech it is common to measure the topography of what was said by (1) *phonemes*: the individual speech sounds that compose a word; (2) *morphemes*: the units with an individual piece of meaning; (3) *lexicon*: the total collection of words that make up a given language; (4) *syntax*: the organization of words, phrases, or clauses in sentences; and (5) *grammar*: the adherence to established conventions of a given language. The formal description of language also involves the classification of words by their grammatical forms or parts of speech, specifically, nouns, verbs, adjectives, adverbs, prepositions, determiners, conjunctions, and interjections (Barry, 1998; Brinton & Brinton, 2010). The traditional classification system is based on vocal response forms; thus certain adjustments are necessary for accommodating nonvocal response forms. For example, sign language involves *cheremes* rather than phonemes (Stokoe, Casterline, & Croneberg, 1965), and selection-based icon systems (e.g., iPad communication devices) have no clear parallel to phonemes.

Function

The second type of classification system involves the different functions that a word may have. For example, the same word (e.g., *book*) can have different meanings (e.g., something to read versus making a hotel reservation). Skinner suggests that the meaning of a word is not to be found in the form of a response, a cognitive processing system, a physical referent, structural regularity, or a person's genetic endowment, but rather in the immediate and historical environmental contingencies that evoke and consequate speaker and listener behavior. The primary purpose of his book *Verbal Behavior* was to identify those environmental contingencies.

Defining Verbal Behavior

Identifying exactly what constitutes language has eluded linguists for centuries. Skinner (1957) argued that what is commonly referred to as communication and language is learned speaker and listener behavior that is acquired, maintained, and extended by the same types of environmental variables that control nonverbal behavior (e.g., reinforcement, stimulus control, motivating operations). Skinner preferred the term *verbal behavior* over the term *language* because verbal behavior emphasizes the individual speakers and listeners rather than the practices of whole verbal communities (e.g., the English *language*).

In defining verbal behavior, Skinner (1957) placed the focus on an analysis of the immediate and historical contingencies that affect speakers and listeners. Accordingly, he defined **verbal behavior** "as behavior reinforced through the mediation of other persons" (p. 2), but these other persons "must be responding in ways which have been conditioned *precisely in order to reinforce the behavior of the speaker"* (p. 225). For example, if a parent asks a child, "Can you let the cat out?" the behavior of the child (listener) is necessary to produce an open door; thus, the receipt of the reinforcer is mediated by the behavior of the child (the listener). But the child must have a training history to successfully mediate reinforcement in this way (e.g., the child understands his[1] parent's words). Nonverbal behavior could produce the same reinforcer by the parent opening the door herself, but the reinforcement for this behavior is not mediated through another person.

Skinner's definition of verbal behavior was not meant to solve the linguist's query regarding the beginnings of "true" language (e.g., Hayes, Barnes-Holmes, & Roche, 2001; Hockett, 1960; Malott, 2003; Normand, 2009; Palmer, 2004). Michael (in Michael & Malott, 2003) noted that "Skinner's definition of verbal behavior . . . is only intended to identify [that] the topic of his consideration in the book on verbal behavior . . . can be narrowed to what is traditionally recognized as the verbal field" (p. 116). By treating language as behavior emitted by speakers and listeners, the focus of the analysis and its applications is on the contingencies responsible for that behavior.[2]

Speaker and Listener

The behaviors of a speaker and a listener are controlled by different but interlocking contingencies. These contingencies and the behaviors they establish and evoke are simple at first, but rapidly grow to enormous complexity. Skinner (1957) provides separate accounts of speaker and listener behavior and terms their interactions *verbal episodes* (p. 38). In a **verbal episode**, a **speaker** emits any type of verbal response in any form (speech, sign language, eye contact), and a **listener** (1) serves as an audience for a speaker, (2) provides reinforcement for a speaker, and (3) responds in specific ways to the speaker's verbal behavior. The roles of speaker and listener alternate in an exchange, and usually involve covert speaker and listener behavior as well. These speaker and listener repertoires will be defined, exemplified, and elaborated on throughout this chapter. However, it is valuable to first make a distinction between verbal responses based on response form and those based on stimulus selection.

Topography-based and Selection-based Verbal Behavior

Verbal behavior is not restricted to spoken words. Any form of operant behavior can acquire a verbal function (e.g., sign language, fingerspelling, icon selection, Braille, facial expressions, intonation). Skinner's broader view of linguistic response

forms expands the options for nonvocal persons in need of augmentative and alternative communication (AAC). In an effort to provide clarity to the issue of verbal response forms, Michael (1985) suggested that a distinction be made between topography-based verbal behavior (e.g., speech, sign language, text) and selection-based verbal behavior (e.g., tablet-based icon-selection programs, Picture Exchange Communication System [PECS], voice output devices).

Michael (1985) states that in **topography-based (TB) verbal behavior** "[t]he *unit of verbal behavior* can be described as an increased strength of a distinguishable topography given some specific controlling variable" (p. 1). This category of verbal behavior can include speech, sign language, writing, fingerspelling, and other response forms in which the listener is affected by a speaker's specific response topography. For example, if a child wants to play with an iPad and signs iPad, that response form informs the listener of the relevant motivation, allowing for the delivery of specific reinforcement (the iPad). Had the child signed "cookie," that topography would inform the listener that a different motivator and consequence are involved.

In **selection-based (SB) verbal behavior**, Michael (1985) states that "the *unit of verbal behavior* can be described as an increased control of the pointing [or selecting] response by a particular stimulus as a result of a different stimulus . . . or establishing operation" (p. 1). For example, the same "iPad," request exemplified above could also occur by selecting the iPad icon from a comparison array. The selected icon also informs the listener of the relevant motivation and allows for the delivery of specific reinforcement. In SB verbal behavior, what is conveyed to the listener is the information on the stimulus selected, as opposed to TB verbal behavior, where the listener's behavior is controlled by the topography of a speaker's response.

Michael's distinction between TB and SB verbal behavior has provided valuable conceptual and practical guidance for understanding a number of important behavioral issues (for a review, see Shafer, 1993). For example, the distinction can facilitate the selection and implementation of AAC systems for children with severe language delays (e.g., Charlop-Christy, Carpenter, Le, LeBlanc, & Kellet, 2002; Frost & Bondy, 1994; McGreevy, Fry, & Cornwall, 2012; Sundberg, 1993a; Tincani, 2004). In addition, research shows that TB systems may produce types of generative learning as observed in stimulus equivalence and relational framing more efficiently than SB systems (Jennings & Miguel, 2017; Lowenkron, 1998; Pérez-González, Salameh, & García-Asenjo, 2018; Polson & Parsons, 2000; C. T. Sundberg & Sundberg, 1990; C. T. Sundberg, Sundberg, & Michael, 2018; Wraikat, C. T. Sundberg, & Michael, 1991).

Behavioral Cusps

When an infant learns to walk, it changes her life forever. Once standing and walking occur, new stimuli, motivators, and reinforcers are available and they immediately begin to evoke and shape new behaviors. Rosales-Ruiz and Baer (1997) termed these significant human developmental changes *behavioral cusps*, and defined them as "any behavior change that brings the organism's behavior into contact with new contingencies that have even more far-reaching consequences" (p. 533). A cusp represents the beginning of a new level of learning for a child. Rosales-Ruiz and Baer (1997) elaborate on their definition by noting, "What makes a behavior change a cusp is that it exposes the individual's repertoire to new environments, especially new reinforcers and punishers, new contingencies, new responses, new stimulus controls and new communities of maintaining or destructive contingencies. When some or all of those events happen the individual's repertoire expands . . . and perhaps that leads to some further cusps" (p. 534).

There are many types of cusps; some are minor, but some are especially important for human development because their far-reaching learning effects can change the course of a child's development. Some of a child's most important early cusps involve motor behavior (e.g., a pincer grasp, crawling, walking), social behavior (e.g., joint attention, adults as conditioned reinforcers), and language acquisition (e.g., imitating, echoing, listening, manding) (Bosch & Fuqua, 2001; Bosch & Hixson, 2004; Greer & Ross; 2008; Hixson, 2004; Hixson, Reynolds, Bradley-Johnson, & Johnson, 2010; Rosales-Ruiz & Baer, 1997; Staats, 1996).

It is important to distinguish between the acquisition of a cusp and the facilitative effects of the cusp itself. In acquisition, the elements of a cusp are gradually established, but the new consequences produced by the cusp only become available when the cusp is achieved. For example, approximations to crawling are gradually shaped, but only when crawling is successful does its far-reaching consequences begin to affect an infant's behavior in a more powerful way (e.g., the automatic reinforcement obtained by crawling to a different room). Cusps can be valuable to a child, especially when they are combined with other cusps and produce even more generative cusps (Greer & Ross, 2008; Hixson et al., 2010; Michael, Palmer, & Sundberg, 2011; Rosales-Ruiz & Baer, 1997). For example, once a child can emit echoic, tact, listener, and self-listener behaviors, new mands and tacts can be acquired incidentally with no formal training or reinforcement history (Horne & Lowe, 1996; Pérez-González, Pastor, & Carnerero, 2014).

The Elementary Verbal Operants and Listener Discriminations

A behavioral analysis of expressive and receptive language involves identification of the antecedents and consequences that control the different types of speaker and listener behavior. Skinner (1957) identified three *separate* sources of antecedent control for speaker behavior: (1) motivating variables, (2) nonverbal discriminative stimuli (S^Ds), and (3) verbal S^Ds. These three sources of antecedent control and their related history of consequences provide the framework for distinguishing between five different types of expressive language (duplic, mand, tact, intraverbal, codic). Michael (1982) termed these basic components of Skinner's taxonomy the **elementary verbal operants**. Receptive language involves the behavior of the listener, and can be behaviorally characterized as mainly nonverbal listener discriminations controlled by verbal S^Ds (e.g., a teacher says, "touch the red balloon" and the child touches the red balloon, not the blue one). However, much of

Figure 18.1 The elementary verbal operants and listener discriminations.

SPEAKER REPERTOIRES (Expressive Language)		
Antecedent Variables	**Behavior**	**Listener-mediated Consequences**
Motivating operation (value of being pushed is high)	*Mand* (saying "push")	Specific reinforcement (being pushed)
Nonverbal S^D (seeing snow)	*Tact* (saying "snow")	Generalized S^r (e.g., praise, approval)
Verbal S^D with formal similarity to the response (hearing "bear")	*Duplic (echoic, motor imitation, copying text)* (saying "bear")	Generalized S^r
Verbal S^D with *no* formal similarity, but with point-to-point correspondence (seeing "dog" written or hearing "dog")	*Codic (textual, taking dictation)* (saying "dog" or spelling "dog")	Generalized S^r
Verbal S^D *without* point-to-point correspondence (hearing "What time?")	*Intraverbal* (saying "10 AM")	Generalized S^r

LISTENER DISCRIMINATION (Receptive Language)		
Antecedent Variables	**Behavior**	**Consequence**
Verbal S^D and a nonverbal S^D ("Give me a pen" along with the pen)	Nonverbal operant behavior (listener hands the speaker a pen)	Generalized S^r

a listener's behavior is also verbal (e.g., covert self-talk) and should be treated as verbal behavior (Schlinger, 2008a). The elementary verbal operants and listener discriminations will be briefly defined and exemplified below and in Figure 18.1, followed by a more in-depth treatment of each type of speaker and listener repertoire.

Mand

When a person asks for something to drink, for directions to a post office, or to be left alone, she is usually manding. Mands allow a speaker to get her needs and wants known to a listener. The **mand** relation is a type of verbal behavior in which the form of a response (words, signs, gestures, icon exchanges, etc.) is under the functional control of a motivating operation (MO) and a history of specific reinforcement (Figure 18.1). For example, a mand can be observed when a child wants to be pushed on a swing and emits the word "push" and an adult then pushes her on the swing. The reinforcement for a mand is specific to the MO; that is, being pushed by adults is what the child wants, not just claps and cheers.

Tact

When a child names her friends in the room while looking at them, she is usually tacting. Tacts involve a speaker's ability to verbally identify aspects of the physical environment. The **tact** relation (Figure 18.1) is a type of verbal behavior in which the form of a response is under the functional control of a nonverbal discriminative stimulus (S^D) and a history of conditioned reinforcement (S^r) (e.g., seeing snow and saying "snow").

Duplic

There are three types of **duplic** relations: echoic, motor imitation, and copying text. They all have the same defining features but have unique topographical features. Repeating spoken words exemplifies the echoic, imitating sign language exemplifies motor imitation, and copying written words exemplifies copying text.

Duplic: Echoic. The **echoic** relation (Figure 18.1) is a type of duplic verbal behavior in which the form of a vocal response is under the functional control of an auditory verbal S^D that has formal similarity between the stimulus and the response product, and has a history of generalized S^r (Figure 18.1). **Formal similarity** occurs when the controlling verbal stimulus and the response product are in the same sense mode (e.g., both are auditory), and the stimulus and the response product resemble each other in the physical sense of resemblance (e.g., hear "bear," say "bear").

Duplic: Motor Imitation. **Motor imitation** (relating to sign language) (Figure 18.1) is a type of duplic verbal behavior in which the form of a motor response is under the functional control of a visual verbal S^D that has formal similarity and a history of generalized S^r (e.g., parent signs "car" and the child signs "car").

Duplic: Copying Text. Writing letters and words that match other letters and words constitutes copying text (fingerspelling and Braille can also be copied). Copying textual stimuli has the same defining features as echoic and motor imitation, except the response form is written, fingerspelled, or Braille punched. **Copying text** (Figure 18.1) is a type of verbal behavior in which the form of a transcriptive response is under the functional control of a textual verbal S^D that has formal similarity and a history of generalized S^r (e.g., writing down an address observed on a website).

Codic

There are two types of **codic** relations: textual and taking dictation. Textual behavior for speech is emitting vocal

responses controlled by written verbal S^Ds, while taking dictation is emitting transcriptive and spelling behaviors controlled by vocal verbal S^Ds. Similar codic relations can occur with sign language and Braille.

Codic: Textual. The **textual** relation (Figure 18.1) is a type of verbal behavior in which the form of a response is under the functional control of a verbal S^D and a history of generalized S^r, but has *no* formal similarity between the stimulus and the response product. However, there is a broader type of match between the verbal stimulus and the verbal response that Skinner terms *point-to-point correspondence*. **Point-to-point correspondence** occurs when the parts of a verbal stimulus correspond to the parts of a verbal response but do not physically match each other (e.g., see "D-O-G," then saying "dog").

Codic: Taking Dictation. When someone gives you a name and you write it down, you convert spoken words to written words (spelling). Although somewhat dated, but nonetheless accurate, Skinner's term **taking dictation** (Figure 18.1) describes this type of verbal behavior, in which the form of a response is under the functional control of a verbal S^D with point-to-point correspondence between the verbal stimulus and the verbal response product, and a history of generalized S^r, but has no formal similarity (e.g., hearing "dog," then typing "D-O-G").

Intraverbal

Answering questions, telling stories, and recalling memories are usually intraverbal behaviors. Our intraverbal repertoires allow us to verbally respond to verbal stimuli we encounter or produce ourselves. The **intraverbal** relation (Figure 18.1) is a type of verbal behavior in which the form of a response is under the functional control of a verbal S^D that does *not* have point-to-point correspondence with the verbal stimulus (e.g., receiving the text "What time?" and texting back "10AM"). There is also a relevant history of generalized S^r involved.

Listener Discriminations

An important component of language involves a listener's ability to discriminate between verbal stimuli emitted by speakers, and to act upon those stimuli, even when the speaker and listener are the same person. Listener discriminations are commonly referred to as receptive language. **Listener discriminations** (Figure 18.1) can be defined as a verbal S^D that evokes a corresponding nonverbal response (or class of responses) due to a history of generalized S^r (e.g., someone asks for a pen and you reach for a pen).

Matching-to-Sample (MTS)

MTS procedures teach a child to attend to and organize the various stimuli that affect his sensory systems (e.g., sights, sounds, tastes). A generalized matching repertoire (e.g., placing novel pictures of various exemplars of a category together) can facilitate the establishment of equivalent and nonequivalent relations between various classes of stimulus events (Chapters 19 and 20). Matching skills are valuable in language instruction for severely language-delayed children and adults. For example, a child who can match pictures of dogs, and discriminate them from other animals, and then learns to call one of the pictures a "dog," may be able to correctly tact novel pictures of a dog, without further training.

Classifying the Verbal Operants

An important component of making use of Skinner's analysis of verbal behavior is the ability to classify a verbal response as a mand, tact, intraverbal, and so on. Classifying the verbal relations allows us to more precisely identify, organize, and account for the specific variables that might control a given utterance. Classification can also assist in identifying fine distinctions between variables that may prove to be of interpretive or clinical value (Palmer, 2016). Figure 18.2 presents a flow chart that can be used to narrow down the sources of control for a sample of verbal behavior. Verbal classification can be accomplished by asking the following questions regarding possible controlling variables for a target response.

1. Does an MO control the response form? If yes, then the response is at least part mand.
2. Does an S^D control the response form? If yes, then:
3. Is the S^D nonverbal? If yes, then the response is at least part tact.
4. Is the S^D verbal? If yes, then:
5. Is there formal similarity between the verbal S^D and the response? If yes, then the response is duplic (echoic, imitative, or copying text). If not, then:
6. Is there point-to-point correspondence between the verbal S^D and the response? If yes, then the response is codic (textual or taking dictation). If not, then the response is at least part intraverbal.

THE VERBAL OPERANTS AND LISTENER BEHAVIOR IN MORE DETAIL

Motivating Operations and the Mand

Motivation is an important source of antecedent control of human behavior. Much of what we do each day is controlled by MOs (Chapter 16). Unlearned MOs, such as those related to food, water, warmth, and sleep, are straightforward and need to be accommodated each day. In contrast, learned MOs, such as those related to money, information, travel, relationships, and so on, are often more complicated in part due to their variability and uniqueness to each individual.

Both types of MOs can evoke nonverbal and verbal behaviors. In addition, their related histories of specific reinforcement can alter the functions of other MOs and S^Ds (Schlinger & Blakely, 1987, 1994). For example, on a hot day when a person is shopping and has not had anything to drink for a while, the value of water might gradually increase to the point where nonverbal behavior is evoked (e.g., looking for a drinking fountain). The history of specific reinforcement related to the MO can also alter the functional properties of other stimuli, such as establishing a drinking fountain as an S^D that immediately evokes approach behavior. If water is located and a sufficient amount is consumed, the value of water and its evocative strength weakens via satiation,

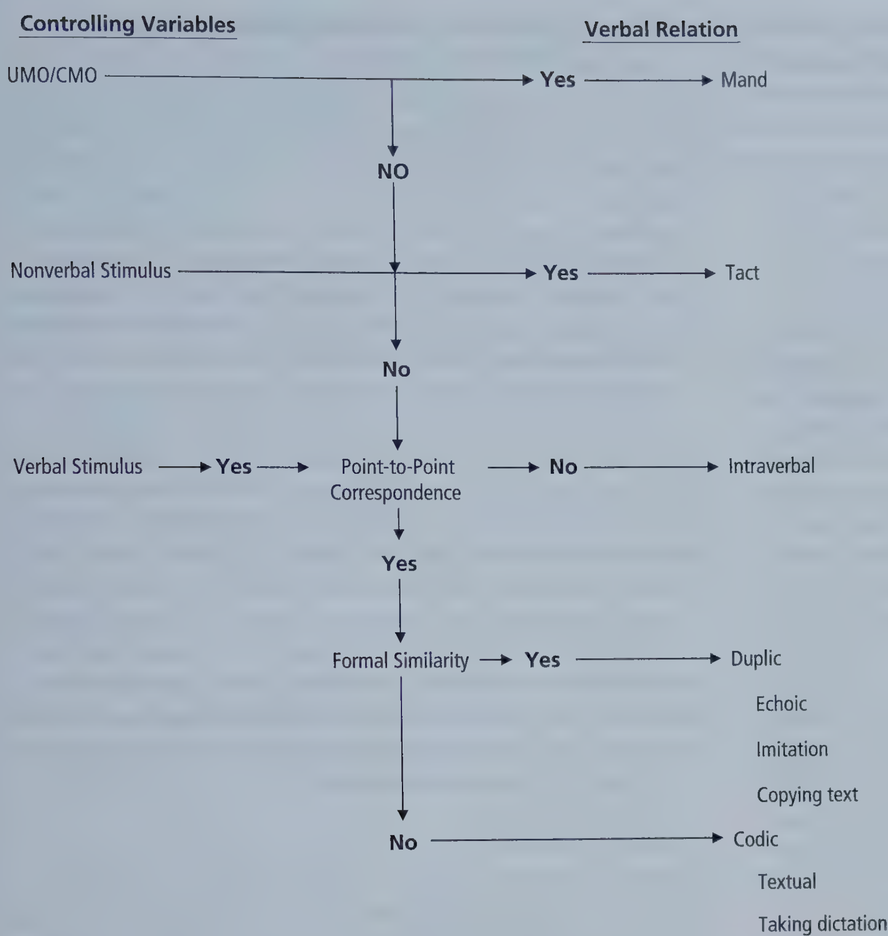


Figure 18.2 Verbal behavior classification chart.

thus an abolishing operation (AO), and the functional properties of the fountain change to an S^{Δ} (one stops searching for drinking fountains).

The same MO (water deprivation) can also evoke verbal behavior (e.g., asking for water), but a listener is necessary to mediate (deliver) the reinforcement, and to complete the verbal episode. In this case, a specific history of reinforcement from listeners alters the functional properties of MOs, as well as the functional properties of potential listeners, establishing some listeners as S^Ds for manding behavior (e.g., a waiter) and others as S^{Δ}. When the mand is successful, it results in the same specific reinforcement and AO achieved through nonverbal behavior that involves obtaining water, with the only difference being how the water was obtained. When MOs evoke a verbal response that produces specific reinforcement mediated by listeners, Skinner (1957) termed this verbal relation a *mand* (Figure 18.1).

Specific Reinforcement and the Mand

The defining consequences for the mand constitute what Skinner (1957) terms *specific reinforcement* (p. 56). Specific reinforcement involves reinforcement that is established by its relation to a specific MO (e.g., water deprivation establishes water as specific reinforcement). When sufficient specific reinforcement is provided, it abolishes the MO. For example, if there is an MO at strength for finding one's car keys, the specific reinforcement is locating the keys. This specific reinforcement also alters the function of any number of contextual stimuli (e.g., drawers become S^Ds for opening and searching behaviors). Finding the keys has multiple effects; it abolishes the MO and simultaneously functions as reinforcement, strengthening the particular behavior that was evoked by the MO (e.g., searching, manding). Specific reinforcement can be contrasted with generalized S^r characteristic of the other verbal operants. Generalized S^r acquires its strengthening effect by its relation to a variety of MOs and relevant reinforcement history, and, as a result, is not dependent on any single MO (Michael, 2004).

Classifying Mand Relations

Classifying the mand relation largely depends on talking precisely about its two defining features: (1) MOs and (2) specific reinforcement. Mands can be classified by identifying the different types of MOs that control manding behavior and by the different ways mands affect the behavior of the listener and produce specific reinforcement. Michael and colleagues (e.g., Laraway, Snycerski, Michael, & Poling, 2003; Michael, 1982, 1993, 2004; Chapter 16) distinguish between two main types of MOs: unconditioned motivating operations (UMOs) and conditioned motivating operations (CMOs). UMOs involve types of motivation that are innate to humans, and do not require a learning history to affect behavior, while CMOs are the direct result of a learning history. In general, a distinction could be made between (1) MOs related to deprivation

and specific reinforcement established by that deprivation, and (2) MOs related to the removal of aversive stimulation and the specific negative reinforcement related to that condition. There are different types of UMOs and CMOs that evoke mands, each of which has different behavioral effects (Chapter 16).

Mands Evoked by Unconditioned Motivating Operations (UMOs)

Michael and Miguel (Chapter 16) identified nine UMOs, six of which are related to deprivation (food, water, sleep, activity, oxygen, and sex) and three of which are related to aversive stimulation (the aversive conditions of becoming too hot, becoming too cold, and an increase in painful stimulation). Each of these MOs (e.g., sleep deprivation) can evoke learned nonverbal behavior (e.g., going to bed) and/or verbal behavior (e.g., a mand), and the specific reinforcement related to the MO can alter the functional effects of other variables as well (e.g., a sofa functions as an S^D evoking an approach response).

Deprivation-based UMOs. Some of an infant's first mands are evoked by unlearned deprivation (Bijou & Baer, 1965; Skinner, 1957). What begins as a human biological process soon becomes learned operant behavior. Food deprivation, for example, may initially elicit crying as a biological (respondent) event, but through the process of reinforcement (i.e., feeding the infant), crying becomes an operant (Bijou & Baer, 1965; Novak & Pelaez, 2004; Schlinger, 1995). Each of the different deprivation-based UMOs may initially evoke crying, but due to the effects of differential specific reinforcement, the response topographies begin to shift (e.g., a hunger cry sounds different from a tired cry). Parents are often able to discriminate between different forms of crying by about the 8th week of an infant's life (Brazelton & Sparrow, 2006). Eye contact can also serve a similar verbal function for an infant in that MOs and specific reinforcement can establish differential looking behavior that acquires a mand function (Isaksen & Holth, 2009).

Aversive Stimuli as UMOs. Three UMOs are related to unlearned aversive stimulation (too hot, too cold, and painful stimulation), and it is their removal that functions as specific negative reinforcement. Michael and Miguel (Chapter 16) explain that "painful stimulation establishes pain reduction as a reinforcer and evokes the (escape) behavior that has achieved such reduction" (p. 381). For example, if a child experiences a burn, the behavior evoked by this aversive UMO can be nonverbal (e.g., turning on a cold water faucet), or it can be verbal (e.g., "Mommy!"). Mands controlled by aversive stimulation also occur early in language development and, like mands related to deprivation, participate in many aspects of human development (e.g., language acquisition, bonding, intellectual development). Identifiable mands of any type are valuable to caregivers in that a recognized response form can help a caregiver identify a child's private MOs and the desired specific reinforcement. Thus, early manding allows reliable access to a child's inner world.

Mands Evoked by Conditioned Motivating Operations (CMOs)

CMOs are more varied than UMOs and can change frequently. Michael and Miguel (Chapter 16) define a CMO as "motivating variables that alter the reinforcing effectiveness of other stimuli, objects, or events, but only as a result of the organism's learning history" (p. 393). For example, a low battery notice on a phone increases the value of a power cord and evokes behavior that involves a history of obtaining that specific reinforcement. Michael and Miguel (Chapter 16) identify three types of CMOs: transitive (CMO-T), reflexive (CMO-R), and surrogate (CMO-S), each of which can evoke manding behavior.

Parts of Speech. Another way to classify some CMO effects is by the parts of speech. For example, a CMO-T related to nouns often involves a stimulus change (e.g., encountering a parking meter) that alters the value of objects (e.g., coins). Verbs often involve movement, and a stimulus change (e.g., being put in a wagon) can increase the value of an event (e.g., being pulled). Similar types of CMO-Ts can be identified for many adjectives, adverbs, prepositions, etc., as well as for verbal information, help, and so on. For those in need of language intervention, there are many ways in which these types of CMO-Ts can be captured or created for instructional purposes.[3]

Multiple Control

Any instance of human behavior is usually a function of multiple variables operating simultaneously. The task of the behavior analyst is to identify the relevant sources of control responsible for a behavior of concern, and Skinner's analysis of multiple control can be of help in this process. Skinner (1957) begins his treatment of **multiple control** by stating, "Two facts emerge from our survey of the basic functional relations in verbal behavior: (1) the strength of a single response may be, and usually is, a function of more than one variable and (2) a single variable usually affects more than one response" (p. 227). Michael et al. (2011) suggested that Skinner's two effects be termed (1) *convergent multiple control* and (2) *divergent multiple control*. In convergent multiple control, a single response is controlled by more than one antecedent variable. For example, a mand is usually controlled by both an MO (e.g., for information) and an S^D (e.g., an audience). In divergent multiple control, a single antecedent variable controls more than one response. For example, the verbal stimulus "Who was Darwin?" could evoke any number of different verbal responses from a listener turned speaker, such as "natural selection," "HMS Beagle," or "the Galapagos Islands."

Convergent Control Involving Multiple MOs. Multiple types of motivation can participate in any verbal episode. MOs can combine and enhance an evocative effect, or compete with each other for the control of behavior (e.g., Michael et al., 2011; Sundberg, 2013). For example, when MOs related to exercise and social interaction combine, they may evoke a mand like "Do you want to go to the gym?" MOs also commonly compete with each other. For example, an MO related to food deprivation and

social embarrassment may compete to evoke either "Yes, I'd like another piece of cake" or "No, thank you." This convergence of antecedent MO variables on behavior often produces emergent effects (both respondent and operant) quite different from the effects of each element of the compound arrangement operating independently.

Convergent Control Involving MOs and S^Ds. MOs typically participate with S^Ds in evoking our day-to-day behavior. As with multiple MOs, combining MOs with S^Ds can produce an additive or competitive effect on behavior. For example, early mand training for nonverbal children consists of combining echoic or imitative prompts with MOs and salient nonverbal stimuli. MOs and S^Ds can also compete with each other. For example, a parent may have strong instructional control over a child's behavior, except when he is playing video games. S^Ds can affect MO control as well, as when a verbal S^D (e.g., a question about sports) disrupts control by a current CMO-R (e.g., studying for an exam).

Mand Control. An MO may be insufficient to control behavior on its own, but when combined with other antecedents, the MO successfully participates in evoking behavior. For example, an MO regarding helping people may not evoke behavior until other stimuli, such as a person falling, are combined with the incipient MO. The weakness of an MO related to helping people could be demonstrated by seeing someone fall and not offering help. Palmer (2016) suggests that there are many circumstances in which the "antecedent is, by itself, insufficient to evoke the relevant response" (p. 100), but when combined with other variables it evokes behavior. In his discussion on intraverbal relations, Palmer suggests the term *intraverbal control* be used to identify this supplementary type of control as it relates to verbal S^Ds. A similar case could be made for *mand control* where an MO by itself is insufficient to evoke behavior, but, when combined with other variables, an evocative effect occurs.

Private Events

Most types of MOs are private; that is, only one person can experience their effects. Skinner (1953) termed these types of covert variables **private events** and stressed that despite their inaccessibility, private events cannot be neglected as causal variables in the interpretation of behavioral events (Skinner, 1945, 1953, 1974). Skinner (1953) explains: "With respect to each individual . . . a small part of the universe is *private*. We need not suppose that events which take place within an organism's skin have special properties for that reason. A private event may be distinguished by its limited accessibility but not, so far as we know, by any special structure or nature" (p. 257).

Many events occur within our bodies and affect how we behave. For example, verbal behavior can easily become covert. A speaker becomes her own listener and responds to her own verbal behavior much in the same way she responds to the verbal behavior of others (Horne & Lowe, 1996). We can and do talk to ourselves and it helps us solve problems, comprehend what others say, edit what we might say, recall past events, entertain ourselves, and so on. Although these private verbal events will always be inaccessible to others, under some circumstances they can be indirectly measured (Palmer, 2011).

Nine Types of Mands Distinguished by Listener Behavior

Skinner suggests that some mands can be classified by the differential mediating behavior of the listener and the reinforcement that strengthens that listener behavior. He offers nine types of mands (Figure 18.3): requests, commands, prayer, questions, advice, warning, permission, offer, and call. For example, in making the distinction between a request and a command, Skinner suggests that a request is emitted to a listener who is motivated to reinforce a speaker (e.g., a waiter). The specific reinforcement provided by the listener is the specific item requested, and reinforcement for the listener is some form of generalized S^r (e.g., "Thanks!"). A command, in contrast, involves an unwilling listener and the verbal stimulus needs to be supplemented with aversive stimuli. The specific reinforcement provided by the listener is compliance with the speaker's command. The listener is reinforced by the removal of the threat—thus a type of negative reinforcement. Skinner provides a similar treatment for the other types of mands he identified (Figure 18.3).

Emergent Mand Relations

Learning to mand is a powerful behavioral cusp in early child development (Bijou & Baer, 1965; Greer & Ross, 2008; Rosales-Ruiz & Baer, 1997; Sundberg & Michael, 2001). A child's mand repertoire typically grows quickly (due to MO and specific reinforcement effects) and often does so with little formal instruction. Once a child has acquired basic echoic, mand, tact, and listener repertoires, the combination of these cusps can produce emergent mands in a number of different ways. For example, once a child acquires mand frames (e.g., "I want_X_") and can tact objects and actions, new tacts can be inserted into the mand frame by direct training, incidental exposure, or relational framing (e.g., Hall & Sundberg, 1987; Horne & Lowe, 1996; Pérez-González, Pastor, & Carnerero, 2014; Rosales & Rehfeldt, 2007). A more advanced and quite generative mand cusp involves mands for information (e.g., "What's that?" "Where's Mommy?"). A child's ability to ask questions plays a major role in vocabulary growth by generating new tacts, echoics, mands, intraverbals, and listener behaviors.

Nonverbal Stimulus Control and the Tact

The distinction between the mand and the tact is based on the distinction between MOs and nonverbal S^Ds as antecedents, and specific and nonspecific reinforcement as consequences. Nonverbal S^Ds, like MOs, can evoke both nonverbal and verbal behaviors. For example, we can respond to rain nonverbally by getting an umbrella, or we can respond to rain verbally by talking about the rain. When the nonverbal properties of the physical environment (e.g., visual or tactile stimulation of rain)

Figure 18.3 Types of mands distinguished by listener behavior.

Mand Type	Speaker's MO	Specific Reinforcement for the Speaker	Reinforcement for Listener Behavior
Request	MO for action on the part of a willing listener	Listener complies	Generalized conditioned reinforcement
Command	MO for action on the part of an unwilling listener	Listener complies	Remove aversive threat
Prayer	MO to generate an emotional disposition	Emits praying behavior	Improves emotional condition
Question	MO for verbal action	Verbal information provided	Generalized conditioned reinforcement
Advice	MO for the listener to enjoy consequences	Listener takes advice	Accesses reinforcement and avoids punishment
Warning	MO to inform the listener of a possible aversive stimulus	Listener following warning	Avoiding aversive stimulus
Offer	MO to get an unwilling listener to participate	Listener participates	Experiences something new
Permission	MO to remove an aversive stimulus	Aversive stimulus removed	Behave without negative ramifications
Call	MO for a listener's attention	Listener emits attending behavior	Reinforcement from additional speaker behavior

evoke a verbal response that is followed by generalized S^r, Skinner (1957) termed this verbal relation a *tact* (Figure 18.1).

Michael (2004) defines nonverbal stimuli as environmental changes (e.g., offset to onset of a stimulus) that affect our receptor systems (e.g., vision, hearing, vestibular). When a particular nonverbal stimulus (e.g., rain) evokes a particular response, and that response is immediately followed by reinforcement, stimulus control develops. Skinner (1957) identified two general types of nonverbal stimulus control that work together as a form of convergent multiple control to evoke verbal behavior: (1) the audience that "controls a large group of responses" (p. 81) and (2) "nothing less than the whole of the physical environment—the world of things and events which a speaker is said to 'talk about'" (p. 81).

Classifying Tact Relations

Classifying tacts can be accomplished by identifying the different types of nonverbal S^Ds that control responses, and by the different ways these verbal responses affect the behavior of the listener. In addition, nonverbal S^Ds commonly interact with other nonverbal S^Ds, verbal S^Ds, and MOs. Several ways to classify various nonverbal S^Ds and tacting will be briefly described.

Different Sensory Receptors

Tacts can be classified by the sensory receptor affected by a nonverbal stimulus. Visual stimuli have functional properties that are different from auditory stimuli, tactile stimuli, olfactory stimuli, and so on. For example, a child may emit the tact "ice cream truck" under the auditory nonverbal stimulus of hearing the truck's music, but not be able to emit the same words under the visual nonverbal stimulus of seeing the truck itself.

Parts of Speech

Tacts can be classified by the parts of speech (e.g., nouns, verbs, adverbs) because a unique type of nonverbal stimulus control is often involved for each part. For example, the nonverbal stimuli that evoke many nouns are typically static objects (e.g., dog, house), while the nonverbal stimuli that evoke verbs often involve movement (e.g., jumping, opening). Many adjectives involve a more complex arrangement of nonverbal stimuli in which the same properties can occur in different objects (e.g., green, large, heavy), and these properties are the source of stimulus control, rather than the object itself. For many prepositions, the objects themselves are irrelevant to the discrimination; rather, their spatial relation to each other is the relevant source of nonverbal control (e.g., next to, on, over). The nonverbal stimulus control involved in other parts of speech (e.g., pronouns, adverbs) can be analyzed in a similar manner.

Multiple Nonverbal Stimuli

Most nonverbal circumstances involve multiple nonverbal stimuli. At a basic level, a single stimulus (e.g., dog) comprises multiple nonverbal stimuli (e.g., a tail, paws, fur). Most encounters with the environment involve multiple nonverbal S^Ds of many types (e.g., objects, actions, relations). For example, the tact "The dog is begging," upon observing a dog, contains multiple nonverbal stimuli, each of which participates in evoking the response. The different parts of speech in the sentence are controlled by a different nonverbal source of stimulus control. The noun "dog" is controlled by the nonverbal object, the verb "begging" is controlled by the dog's actions, and the verb "is" is a relational autoclitic informing the listener of the coordinated relation between the dog (noun) and begging (verb). When a sentence contains other parts of speech (e.g., adjectives,

pronouns) additional nonverbal S^Ds participate in evoking a response. Thus, what may appear to be a single utterance with a single "meaning" is actually multiple responses controlled by multiple variables.

Convergent Multiple Control Involving Nonverbal S^Ds, Verbal S^Ds, and MOs

Nonverbal S^Ds commonly participate in verbal episodes with other types of antecedent variables. For example, nonverbal and verbal S^Ds are involved when a child is asked, "What shape is this paper?" Saying "rectangle" is multiply controlled by the actual (nonverbal) shape of the paper and by the (verbal) spoken word "shape." If the question was "What *color* is this paper? a different response would be evoked. This combination of controlling variables makes the response part tact and part intraverbal, and both sources of control are required to be correct. Nonverbal S^Ds can also share control with MOs. For example, an MO related to a cut finger and the nonverbal S^D of a Band-Aid box evoke selection behavior.

Tact Control

Multiple control can also play a role when the nonverbal antecedent is, by itself, insufficient to evoke the relevant response (Palmer, 2016). For example, if a person cannot correctly remember the name of someone she is talking with, supplemental verbal stimuli (e.g., discussion of a shared work experience) may help to evoke recall. Extending Palmer's (2016) concept of intraverbal control to situations in which the nonverbal stimulus is insufficient to evoke the relevant response can be termed *tact control*, suggesting that the nonverbal stimulus is not the primary source of control, but rather a participant in a multiple configuration of antecedent variables.

Private Nonverbal Stimulus Control

Nonverbal stimuli can arise within the skin of a speaker and evoke verbal behavior in the form of a tact. For example, if one tightens her stomach muscles, this response produces a nonverbal response product that can be tacted through the kinesthetic receptors. That is, one can feel her muscles move and tact this movement, although no one else can. Visual imagining can also provide nonverbal S^Ds that can be tacted (e.g., Horne & Lowe, 1996; Miguel, 2018; Skinner, 1974) and thus provide a sometimes valuable supplemental source of stimulus control for behavior for intraverbal behavior. For example, Kisamore, Carr, and Leblanc (2011) demonstrated how intraverbal behavior could be improved by teaching children to emit visual imagining behavior (e.g., imagining a farm when asked to name some farm animals). Visual imagining can also participate in bidirectional naming relations described in a later section (Horne & Lowe, 1996; Miguel, 2018).

A child first learns to tact her own overt behavior (e.g., "I'm dancing") before learning to tact her own covert behavior (e.g., "I feel tired"). A verbal community establishes self-awareness in several ways (Skinner, 1953, 1974). For example, observable overt collateral behavior can inform the community of a private stimulus (e.g., a child limping), and these circumstances can be used to teach the child to tact the private event for herself. Another way a child learns a tact controlled by private nonverbal stimulation is when the private stimuli are accompanied by observable external stimuli (e.g., blood from a cut). The adults can tact the public stimuli (as well as collateral behavior) that are highly correlated with the private stimuli, and provide and teach an appropriate response form (e.g., "ouch") when the private stimulus is present, thus allowing for echoic-to-tact transfer. Private events can become quite complicated when emotions, MOs, verbal S^Ds, nonverbal S^Ds, and other variables interact with each other and the current context and reinforcement contingencies.

Tact Extension

Once a tact relation has been established, the tact response can occur under novel stimulus conditions through the process of stimulus generalization. Skinner (1957) identifies four different levels of generalization based on the degree to which a novel stimulus shares the relevant or defining features of the original stimulus. The four types of **tact extension** are generic, metaphorical, metonymical, and solecistic. In a generic tact extension, the novel stimulus shares all of the relevant or defining features of the original stimulus. For example, if a child learns to tact "tree" in the presence of a maple tree, when shown an oak tree the child correctly tacts it as "tree." When the child encounters a novel nonverbal stimulus that has some, but not all, of the defining features of tree (e.g., a bush) this antecedent may also evoke the response "tree," and exemplifies a metaphorical tact extension.

Another type of generalization to novel stimuli is a metonymical tact extension, in which a novel stimulus does not share any of the defining features of the original stimulus, but some irrelevant feature acquires stimulus control. For example, a child may have a tendency to call a swing a "tree" if there was a swing on the tree, and it was the swing that acquired stimulus control over the response "tree" due to a history of reinforcement. The fourth type of novel tacting, termed *solecistic extension*, may occur when the property that acquires control has no clear connection to the defining properties. For example, there appears to be no connection between tacting the nonverbal stimulus shoe as a "tree," although these types of solecistic extensions can occur.

Tact Abstraction

Tact abstraction involves learning to tact a property (e.g., color, size, shape) regardless of the object itself. For example, the color red can appear as part of a wide range of nonverbal stimuli (e.g., red cars, red clothing, red fruit). Several of the parts of speech involve abstractions, such as adverbs (e.g., fast, slow) and prepositions (e.g., near, above), where the properties of movement and location, respectively, must acquire control over behavior regardless of the context in which that property appears.

Reinforcement Controlling the Tact

In addition to the different types of antecedent variables that distinguish the mand from the tact, different types of reinforcement define each operant as well (Braam & Sundberg, 1991;

Saunders & Sailor, 1979; Stafford, Sundberg, & Braam, 1988). The mand relation involves specific reinforcement, while the tact relation involves generalized S^r. Skinner (1957) identified five different types of generalized S^r for tacting provided by listeners; three are delivered by a second person serving as a listener: (1) educational reinforcement, (2) escape from or avoidance of aversive stimulation, (3) extending the listener's contact with the environment; and two are self-delivered by the speaker acting as her own listener: (4) automatic reinforcement, and (5) a history of contiguous or correlated usage.

Educational Reinforcement

A child's ability to correctly identify objects and events in her physical environment is an important aspect of language development, and parents and other caregivers take special effort to establish this behavior. Skinner (1957) suggests that the type of generalized S^r involved can be identified as educational reinforcement, and he defined it as "reinforcement supplied primarily because it establishes and maintains a particular form of behavior in the speaker" (p. 84). For example, teaching a child to tact body parts involves antecedents, behavior, and consequences (a discrete trial), with differential reinforcement explicitly provided for purposes of instruction. It could be considered contrived reinforcement contingencies; nonetheless, these contingencies are valuable for establishing a child's early tacting vocabulary or for repairing a child's impaired tact repertoire.

Escape from or Avoidance of Aversive Stimulation

The removal of something that is aversive to a person can function as a type of negative reinforcement, thus strengthening behavior. If a threatening tone of voice accompanies a nonverbal stimulus (e.g., a harsh "What is it?" in a teaching trial), it might be that the removal of the aversive stimulus is a more powerful consequence than social praise (e.g., Carbone, Morgenstern, Zecchin-Tirri, & Kolberg, 2008).

Extending the Listener's Contact with the Environment

The main reinforcement for tacting is the action that listeners take as a function of the speaker's tact. Michael (1991) explains this effect in the following manner: "As a result of the speaker's tact, the listener is able to do something that he could not have done without the speaker's tact, and for this reason provides generalized S^r" (p. 35). For example, if a driver (as a listener) cannot see an oncoming car, a passenger's (as a speaker) tact "There's a car coming" extends the driver's contact with the physical environment, and allows him to behave more effectively. This reinforcement is not contrived and produces effective listener behavior, and the listener provides reinforcement for that reason (e.g., "Thanks").

Automatic Reinforcement

Environmental contingencies can affect behavior without the immediate involvement of other people. For example, the reinforcement obtained from swimming may emanate from the behavior itself—that is, it feels good. Skinner (1957) used the term "automatic" to identify circumstances in which a behavior is evoked, shaped, maintained, or weakened by environmental variables occurring without direct manipulation by other people. Vaughan & Michael (1982) note that Skinner's usage of *automatic* was simply to counteract "any tendency to restrict the concept of reinforcement to those occasions upon which it has been deliberately arranged by another person or group" (p. 218). Skinner (1957) also discussed automatic punishment (p. 375), automatic extinction (p. 164), automatic consequences (p. 442), automatic shaping (p. 58), automatic MO control (p. 220), and automatic stimulus control (p. 416). In short, all the basic principles of behavior can operate on a person automatically, and all **automatic contingencies** have the same defining properties as non-automatic contingences. Thus, antecedents and consequences that are not deliberately arranged can affect behavior, often in a manner more efficient than directly manipulated contingencies (Palmer, 1996; Skinner, 1968).

Among the various types of automatic contingencies, automatic reinforcement has received the most attention from behavioral researchers. Automatic reinforcement can occur in two ways: (1) where the reinforcement emanates from the response product of the behavior (e.g., infant babbling, singing songs, mannerisms, emitting a clever phrase) and (2) where the reinforcement is provided by the behavior's reinforcing effects emanating from the physical environment (e.g., turning on a light, finding a missing phone). Both types of automatic reinforcement can produce an increase in the behavior they follow without involvement from another person.

The automatic reinforcement concept has many applications to language instruction and problem behaviors, along with a growing body of empirical research supporting its use (e.g., Chapter 27; Shillingsburg, Hollander, Yosick, Bowen, & Muskat, 2015; Stock, Schulze, & Mirenda, 2008; Sundberg, Michael, Partington, & Sundberg, 1996; van Harran, 2015; Vollmer, 1994). In addition, Donahoe and Palmer (1994) and Palmer (1996, 1998) describe how automatic reinforcement participates in the development of more complex aspects of verbal behavior. For example, these authors suggest that a child's grammar produces automatic reinforcement when it sounds like grammar emitted by others in their verbal community, but is automatically punished when it sounds odd or unusual. Palmer (1996) described this effect as achieving *parity* with others. Several studies have demonstrated how automatic reinforcement participates in establishing and refining grammatical and syntactical skills (e.g., Østvik, Eikeseth, & Klintwall, 2012; Wright, 2006).

A History of Contiguous or Correlated Usage

Each speaker has a unique and private history of reinforcement relevant to her own verbal behavior. This history is brought to bear on current environmental contingencies, and the two sets of variables interact. For example, if a friend is talking about a recent bike trip to Nova Scotia, and you as a listener have also biked around Nova Scotia, a number of covert verbal behaviors and nonverbal imagery may come to strength. The response products of these covert verbal behaviors prepare you for your upcoming speaker role. "Nova Scotia" as a verbal S^D may covertly evoke "Halifax" or "Immigration Station" as a self-intraverbal. A covert visual image of flowers along the road may also function as a nonverbal S^D that evokes "lupines" as a

self-tact. These effects occur because of a history of reinforcement in which these different stimuli occurred together either contiguously or simply correlated (Palmer, 2016). Skinner terms this effect *contiguous usage*, noting, "In general there is an advantage if responses appropriate to a current situation are strong" (p. 86). These responses are self-strengthened by the automatic reinforcement they produce. In addition, their occurrences evoke speaker behavior that increases the probability that the listener will provide generalized S^r (e.g., "I enjoyed the lupines too").

Generative Learning

Typically developing children demonstrate an explosion of language skills between the ages of 2 and 3 (Brazelton & Sparrow, 2006; Greer & Ross, 2008; Hart & Risley, 1995). During that period, a child begins to emit and comprehend new words daily, and vocabulary size can grow from 100 to 1000 words. A child may combine and expand on these new words in novel ways and in novel contexts, often without direct training or a history of reinforcement (e.g., a toddler who shocks an audience with a swearword). Young children seem to learn new words on their own. However, some children, especially those with autism, struggle with or fail to make this leap in language acquisition.

Some emergent verbal behavior can be explained by stimulus and response generalization, but others involve a process of generalization that goes under the rubric of generative learning. **Generative learning** involves a behavioral effect where previously acquired skills enable or accelerate the acquisition of other skills, without dependence on direct teaching or a history of reinforcement (Alessi, 1987; Horne & Lowe, 1996; Staats, 1996; Stewart, McElwee, & Ming, 2013). There are several productive lines of research on generative learning, each of which has contributed to our understanding of this important behavioral effect (e.g., Becker, 1986; Engelmann & Carnine, 1982; Goldstein, 1983; Greer & Ross, 2008; Hayes et al., 2001; Hixson, 2004; Horne & Lowe, 1996; Johnson & Layng, 1994; Koegel & Koegel, 1988; Lowenkron, 1998; Malott, 2003; Rosales-Ruiz, & Baer, 1997; Sidman, 1994; Staats, 1996). The concept of behavioral cusps (Rosales-Ruiz & Baer, 1997) can help organize the topic of generative learning, especially as it applies to early language acquisition (Greer & Ross, 2008). In addition, the work of Staats (1996) on cumulative-hierarchical learning can help to provide a conceptual foundation for generative learning (Hixson et al., 2010). The following section will describe several ways that generative learning can produce emergent tact relations.

Emergent Tact Relations

Stimulus Equivalence

Sidman (1971) and Sidman & Cresson (1973) showed that combining matching-to-sample (MTS), echoic, and listener skills generated emergent MTS, tact, and textual relations. Sidman termed this effect *stimulus equivalence* (Chapter 19). These researchers showed that equivalence can be established by using conditional discrimination procedures to directly teach a single AB relation between a word (stimulus A) and an object (stimulus B), then teach a BC relation between the same object (stimulus B) and a different stimulus (stimulus C), and then test for the emergence of the untrained relations (i.e., BA/CB symmetry tests and AC/CA transitivity tests). When emergence occurs with both symmetry and transitivity, equivalence has occurred (Sidman & Tailby, 1982). Pilgrim (Chapter 19) describes the emergent tact and textual responses in Sidman's early work as follows: "The exciting result was that these boys not only named the pictures correctly, but also the written words; that is they read the words orally where they had not been able to do so prior to the teaching phases" (p. 456). Since Sidman's initial work, equivalence procedures have proved an effective method for producing a host of generative learning effects (Chapter 19).

Recombinative Generalization

Goldstein (1983) defined recombinative generalization as "differential responding to novel combinations of stimulus components that have been included previously in other stimulus contexts" (p. 280). For example, if a child is taught to tact "green circle" and "red square" and then without direct training the child is able to correctly tact "green square" and "red circle," then recombinative generalization has occurred. Matrix training is an example of a commonly used procedure that can produce recombinative generalization, and studies have demonstrated its effectiveness in generating emergent tact relations (e.g., Axe & Sainato, 2010; Frampton, Wymer, Hansen, & Shillingsburg 2016; Kohler & Malott, 2014; Pauwels, Ahearn, & Cohen, 2015).

Common Bidirectional Naming (C-BiN)

Horne & Lowe (1996) suggested that one of the most far-reaching achievements in early language acquisition occurs when the echoic, tact, listener, and self-listener skills combine to establish a new higher order verbal cusp they term common **bidirectional naming**. Once a child can easily echo new words, and has acquired the basic prerequisite repertoires for C-BiN, a new word acquired as listener can immediately generate speaker behavior (a tact) without further training; or if the word is first acquired as a tact, it can immediately generate a listener relation without further training (consistent with symmetry and mutual entailment). A reinforcement history of naming experiences (multiple exemplar instruction) and self-echoic mediation establishes bidirectional relations between words and objects or events, and fuses the speaker and listener repertoires (Greer & Ross, 2008; Horne & Lowe, 1996; Miguel, 2016, 2018). The naming cusp is far-reaching, so much so that Greer and Longano (2010) assert the following: "Naming appears to be the source of the explosion in language development" (p. 73). In addition, C-BiN teaching procedures have shown effectiveness in producing generative learning for children with disabilities who have failed to acquire this important form of learning (e.g., Fiorile & Greer, 2007; Lee, Miguel, Darcey, & Jennings, 2015).

Joint Control

Lowenkron (1984, 1988, 1989) demonstrated that emergent speaker and listener behavior could be generated by a type of multiple control he termed *joint control*. Joint control also involves the combination of verbal skills (e.g., echoic and tact repertoires) and occurs when two separately established

antecedents that evoke the same response topography arise simultaneously. For example, when searching for a specific reference from a list of references, a self-echoic (e.g., "Guttman and Kalish") is rehearsed (overtly or covertly) until a textual stimulus evokes the same response form (e.g., reading "Guttman and Kalish"). The physical similarity of the two response forms generates a nonverbal S^D regarding that similarity and evokes an emergent autoclitic tact (e.g., "that's it"). The response product of the autoclitic tact, along with the textual stimulus, sets up the conditional discrimination that evokes the emergent listener behavior (e.g., copying and pasting the reference). Once joint control is acquired as a generalized repertoire (a higher-order cusp), its generative effects can allow for, or accelerate, a host of complex speaker and listener behaviors (e.g., Causin, Albert, Carbone, & Sweeney-Kerwin, 2011; Lowenkron, 1998, 2006; C. T. Sundberg et al., 2018; Tu, 2006).

Relational Frame Theory (RFT)

Hayes et al. (2001) extended Sidman's equivalence framework to nonequivalent relations such as comparison, opposition, and difference (Chapter 20). Generative learning opportunities that can produce emergent tacts are abundant when equivalent and nonequivalent relations interact in novel contexts. Once acquired as a generalized operant, relational responding functions as a higher-order cusp capable of generating a number of far-reaching derived effects (e.g., Dougher, Augustson, Markham, Greenway, & Wulfert, 1994; Hayes et al., 2001; Stewart et al., 2013). In RFT, framing is established through a history of multiple exemplar training and combined in various ways dependent on the current context. This history and the current contingencies provide the foundation for the establishment of derived relational responding.

A defining feature of derived relational responding is the transformation of stimulus functions, which includes all types of transfer (e.g., mutual entailment, combinatory entailment, transformation of function) (Dymond & Rehfeldt, 2000; Hayes et al., 2001). Unique to RFT is the treatment of the transformation of psychological function (e.g., Dougher et al., 1994; Dymond & Barnes, 1994). In this type of transformation, "[i]f one stimulus in a mutual or combinatorially entailed relation is given a direct psychological meaning, then the remaining stimuli may acquire this psychological function. For example, if 'lemons' is in an equivalence class with actual lemons in a context that selects taste as the relevant function, talking of lemons can be associated with salivation or puckering" (Hayes, Gifford, & Ruckstahl, 1998, p. 289). For individuals with language delays, the transformation of function and related relational framing activities can provide teaching tools for establishing generative networks of stimulus relations that can be emitted, expanded, and merged under a variety of contexts (e.g., Dixon, 2016). To learn more about derived relational responding, see Chapter 20.

Verbal Behavior Controlled by Verbal S^Ds

Verbal S^Ds constitute the third type of antecedent control for language, joining MOs and nonverbal S^Ds. as sources of control for both speaker and listener behavior. We commonly respond to the words of others, and to our own words as a speaker or a listener. We have conversations, surf the Internet, read, write, and so on. Skinner (1957) defines a verbal stimulus as "the product of earlier verbal behavior" (p. 65). That is, verbal responses produce a response product. For example, vocal speech produces auditory stimuli, sign language produces visual stimuli, and Braille dots produce tactile stimuli. These response products can in turn have a discriminative or motivative function evoking the behavior of listeners, including the speaker as her own listener.

Verbal S^Ds have the same causal status as nonverbal S^Ds in that both types of S^Ds acquire discriminative control over behavior through a history of differential reinforcement. Verbal S^Ds have unique effects and characteristics not found with nonverbal S^Ds (Sundberg, 2016a). For example, verbal S^Ds can be portable, private, and free from environmental support. We can talk about being in Italy without being in Italy, recite the numbers to open a combination lock, solve problems, or think through an important decision, all in the absence of the relevant physical stimuli. The verbal response products of this talk may function as automatic S^r, as well as S^Ds or MOs evoking other behaviors. Nonverbal S^Ds, by contrast, are restricted to the immediate physical environment (with the exception of imagining), and are far less portable or available to a speaker across environments.

Six types of elementary verbal operants are controlled by verbal S^Ds: echoic, motor imitation (relating to sign language), copying text, textual, taking dictation, and intraverbal (Figure 18.1). All of these verbal operants, except the intraverbal, can be classified as duplic or codic.

The Duplic Relations: Echoic, Motor Imitation, and Copying Text

Duplics are the least complex of the verbal operants but play a significant role in linguistic and academic development. This category of verbal operants involves verbally copying verbal stimuli. Verbal stimuli can occur in many forms (e.g., spoken words, written words, gestures, sign language, fingerspelling, Braille dots). These can be organized into three distinguishable types: echoic (auditory match), motor imitation as it relates to sign language (visual match), and copying text (e.g., copying configurations of written letters, fingerspelling, or Braille dots).

Echoic. A child's ability to echo the phonemes and words of others is essential for learning to identify objects and actions. In the presence of a duck a parent might say, "That's a duck, can you say duck?" If the child can respond "duck," then the parent says "Right!" Eventually, the child learns to name a duck without the echoic prompt. Echoic behavior also plays a critical role in more complex forms of verbal behavior, such as incidental verbal learning, derived relational responding, joint control, bidirectional naming, and problem solving (e.g., Hayes et al., 2001; Horne & Lowe, 1996; Lowenkron, 2006; Palmer, 2012; Schlinger, 2008a, 2008b).

Motor Imitation. Some types of motor imitation constitute nonverbal behavior (e.g., following others to get in a line), while others can have the same functional properties as echoic

relations. Motor imitation is a valuable tool for teaching sign language to persons who are deaf, or hearing but unable to produce speech. For example, a child may learn to imitate the sign for popcorn while popping corn, then later emit the sign appropriately without an imitative prompt.

Copying Text. Textual verbal stimuli can be written, fingerspelled, or punched in Braille dots. Learning to copy these stimuli can facilitate the development of codic verbal behavior (e.g., learning to read, write, and spell). For example, a child often first learns to make letters by tracing and copying them, which can facilitate the process of independently writing, spelling, and reading.

The Codic Relations: Textual and Taking Dictation

Codic relations are more complicated than duplic relations because of the lack of formal similarity between the verbal S^D and the verbal response. Michael (1982) notes that the term codic "is meant to suggest the kind of relation seen in a formal code, where one stimulus is said to stand for another stimulus that it does not resemble" (p. 2). In codic relations, additional speaker behaviors are required to learn to code and convert a textual stimulus into a vocal stimulus, and vice versa.

Textual. Textual behavior is reading without an implication that the reader understands what is being read. Understanding what is read is typically identified as reading comprehension and involves a number of functionally separate verbal and nonverbal operants (e.g., intraverbal behavior, manding, tacting, listener discriminations). For example, saying "Chicago" upon seeing the written word "Chicago" is usually textual behavior. But understanding that Chicago is a city on Lake Michigan is not textual behavior (it is more likely intraverbal behavior). Textual behavior can occur with sign language as well. The Stokoe Notation System was designed as a form of written sign language (although not widely used). It has 55 written cheremes that are analogous to the 42 vocal phonemes of English (Stokoe et al., 1965). An arrangement of written cheremes can be read with signs in a manner similar to speech because they have point-to-point correspondence to a specific sign.

Taking Dictation. Taking dictation consists of writing and spelling words that are spoken. The multiple repertoires involved in learning this code include not only the manual production of letters (written, typed, fingerspelled, or punched) but also the spelling of words. As with textual, the code aspect of this repertoire usually requires additional instruction. Taking dictation can occur with sign language as well. This relation involves seeing a sign and writing the corresponding cheremes.

The Intraverbal

The intraverbal relation is different from the other types of verbal behavior evoked by verbal S^Ds in that *no* point-to-point correspondence exists between the verbal stimulus and the verbal response (Figure 18.1). Young children typically emit intraverbal responses in the form of singing songs, telling stories, describing activities, and so on. At a more advanced level, intraverbal behavior is a key component of many important aspects of human behavior (e.g., academics, social interaction, entertainment). As with the other verbal operants, the form of an intraverbal response can involve speech, signs, written words, and so on. And as with the mand and tact, to understand the intraverbal, one must understand its controlling variables. At least four different types of verbal discriminations can be involved in intraverbal relations (Figure 18.4): (1) simple verbal, (2) compound verbal, (3) verbal conditional, and (4) verbal function-altering.

Classifying Verbal Discriminations

Intraverbal relations can be classified by identifying the type of verbal stimulus control that evokes the intraverbal response. In general, each of the verbal discriminations presented below involve increasingly more complex types of verbal stimulus control.

Simple Verbal Discrimination

The first type of intraverbal behavior acquired by typically developing children usually involves **simple verbal discriminations**, where a single-component word or phrase evokes a nonmatching verbal response. For example, upon hearing "Ready, set . . . ," a child says "go," or hearing "Buzz . . . " a child says "lightyear." These early intraverbal responses may be initially multiply controlled by MOs and other supplementary variables (e.g., echoics, tacts), but soon intraverbal stimulus control is established.

Compound Verbal Discrimination

A **compound verbal discrimination** involves convergent multiple control where two or more verbal S^Ds collectively evoke a response. However, the response evoked by the compound stimulus is different from the response evoked by each verbal S^D when presented in isolation (Eikeseth & Smith, 2013). For example, hearing "yellow" might intraverbally evoke any number of responses (e.g., "color," "school bus"), and "fruit" might also evoke any number of responses (e.g., "apple," "orange"), but when combined as "yellow fruit" they evoke a more specific response (e.g., "banana," "lemon"). In a compound verbal discrimination, the individual words do not change the function of each other; they are parts of a larger stimulus configuration, similar to the way individual nonverbal stimuli (e.g., tail, paws) make up objects (e.g., a dog). In the next type of verbal stimulus control presented, the individual words in the antecedent do affect each other.

Verbal Conditional Discrimination (VC^D)

Verbal stimulus control becomes more complicated when verbal stimuli interact with each other in the same antecedent configuration. For example, a conditional discrimination can arise between verbal stimuli in an utterance and more behavior is required to emit intraverbal responses (i.e., questions get harder). A **verbal conditional discrimination** occurs when one verbal stimulus alters the evocative effects of another verbal stimulus

Figure 18.4 Four types of verbal discriminations.

Type of Discrimination	Definition	Example
Simple	A single-component verbal stimulus that evokes a response.	*Speaker*: Saying "meow" after hearing "A kitty says . . ." *Listener*: When asked to "jump" a child emits jumping behavior.
Compound	A verbal stimulus that involves two or more S^Ds that each independently evoke behaviors, hut when they both occur in the same antecedent configuration, a different S^D is generated.	*Speaker:* Saying "blue" after hearing "Red, white, and . . ." *Listener:* When asked to "clap fast," and "clap slow," "walk fast," and "walk slow" the corresponding nonverbal behavior is emitted.
Verbal Conditional	A verbal stimulus that alters the evocative effects of another verbal stimulus in the same antecedent configuration.	*Speaker:* Saying "spoon" and "soap," respectively, when asked, "What do you eat with?" and "What do you wash with?" *Listener:* Pointing to spoon and soap when asked the same questions presented above.
Verbal Function-altering	A verbal stimulus that alters the function of other stimuli or motivating operations (Schlinger & Blakely, 1994).	*Speaker:* Singing at the right time after hearing "When I call your name, sing your part." *Listener:* After hearing "When the doorbell rings, get your father," the correct behavior is emitted at the right time.

within the same antecedent configuration, and that stimulus evokes the response (Axe, 2008; Catania, 2013; Devine, Carp, Hiett, & Petursdottir, 2016; Eikeseth & Smith, 2013; Kisamore, Karsten, & Mann, 2016; Michael et al., 2011; Sundberg, 2016a; Sundberg & C. A. Sundberg, 2011). For example, in the verbal antecedent "Who is not going to dinner?" the reinforcement history related to the conditional word "not" alters the evocative effects of the phrase "going to dinner," changing it to a different verbal S^D that evokes a corresponding intraverbal response. A VC^D effect can be demonstrated by showing a change in the evocative effect when the verbal stimulus is altered to "Who is going to dinner?"

Verbal Function-altering Effect

Verbal stimuli can alter the functional effects of immediate or future S^Ds and MOs and, accordingly, change a listener's behavior. The immediate effects involve straightforward operant and respondent conditioning, but the changes in future behavior are more complex. For example, being told "The bridge is out, turn left at the 7-Eleven and there will be another one in 5 miles" can alter the functional effects of stimuli encountered in the future and evoke verbal and nonverbal behavior at that time (e.g., tacting the 7-Eleven, turning left). Skinner (1957) described this type of effect as *conditioning the behavior of the listener* (p. 357) through instruction. Schlinger and Blakely (1994) and Schlinger (2008b) suggested the term **verbal function-altering** be used to describe the effect that occurs when verbal stimuli immediately condition the behavior of the listener and produce relatively permanent changes in the behavioral function of other stimuli and MOs occurring immediately or later in time The future behavior of a listener can involve any speaker, listener, or respondent relation. Schlinger and Blakely (1987, 1994) suggested that the use of Skinner's (1969) term *rule-governed behavior* be reserved for these types of verbal function-altering effects.

The conditioning of a speaker's future intraverbal behavior would first involve the initial listener conditioning and then an encounter with the newly established verbal stimulus. For example, if a peer says to a teenage friend, "When your mother asks where did you go, tell her the library." Upon encountering the mother's question later, it evokes "the library" as instructed. Under other circumstances, hearing "Where were you?" would have different evocative effects (e.g., "track practice"). However, it is important to note that "this does all not happen in the naïve speaker or listener; it is the end result of a long process of verbal conditioning" (Skinner, 1957, p. 360).

Convergent Multiple Control Involving Verbal S^Ds, Nonverbal S^Ds, and MOs

Verbal S^Ds not only combine with each other as a type of multiple control but also commonly participate with nonverbal S^Ds, MOs, and respondent relations in the control of behavior. As with other examples of convergent multiple control, the effect of combining multiple variables can be additive or decremental. On the one hand, verbal S^Ds can provide a welcome addition to nonverbal stimuli, such as the information provided by a tour guide while viewing an historical site. On the other hand, verbal S^Ds can weaken, block, or distort control by other nonverbal stimuli, verbal stimuli, and MOs. For example, when someone talks during a movie scene the nonverbal action or dialog might be missed.

Intraverbal Control

A distinction can be made between verbal S^Ds that depend on multiple control with nonverbal S^Ds and MOs, and those that do not. Those that do not require multiple control, and have the relevant reinforcement history, can be classified as an intraverbal operant (e.g., upon hearing "What's 6 × 8" a child immediately says "48"). However, some verbal stimuli only supplement other more critical antecedents, but nonetheless play a causal role in evoking behavior. Palmer (2016) recommends that "In cases in which the verbal antecedent is, by itself, insufficient to evoke the relevant response, we should speak of intraverbal control, usually as one of a number of concurrent controlling variables" (p. 99). For example, a common practice in tact training for children with language delays is to hold up an object and say, "What is it?" The verbal stimulus "What is it?" by itself is insufficient to evoke the correct response, but, along with the presented object, it can be an important supplemental S^D that helps to evoke a response from a child learning tacts.

Private Verbal Stimulus Control

Verbal stimuli, like MOs, are free from the support of the physical environment; thus they can be portable and private. We can and do talk to ourselves, often to our benefit (e.g., self-editing, problem solving, thinking) (see Box 18.1). Our own covert verbal behavior provides an abundant supply of verbal stimuli (sometimes too much). These verbal stimuli are no different in function than overt verbal stimuli, except only one person can experience the S^D or MO effect. Covert verbal stimuli can combine with other variables, including overt variables. For example, while traveling through an airport and looking for the right terminal and gate, passengers may engage in a series of overt or covert verbal behavior that assists them in navigating various transitions. This verbal behavior can involve self-echoics (e.g., "Gate B31"), self-intraverbals (e.g., "B31 is in a different terminal"), self-textual (e.g., reading the gate signs), and joint control, all of which can participate as causal variables that ultimately evoke going to the right gate. Covert verbal S^Ds have received little experimental attention due to the privacy issue; nonetheless, they can play an important role in the interpretation of complex behavior (Moore 2008; Palmer, 2011; Schlinger, 2008a; Skinner, 1974; C. T. Sundberg et al., 2018).

Reinforcement Controlling the Duplic, Codic, and Intraverbal Relations

Generalized S^r provided by listeners (e.g., praise, approval) establishes verbal stimulus control over verbal behavior, in the same way it does for nonverbal stimulus control. Four of the five types of reinforcement identified for the tact relation apply to the duplic, codic, and intraverbal relations as well: (1) educational reinforcement, (2) termination of aversive stimuli, (3) automatic reinforcement, and (4) a history of contiguous or correlated usage. The form of reinforcement that does not apply was identified as reinforcement from extending the listener's contact with the physical environment, and that form is unique to the nonverbal stimulus control. However, Skinner identifies an additional type of generalized S^r unique to verbal stimulus control; this type can be termed *facilitative reinforcement.*

Facilitative Reinforcement. Skinner (1957) states that verbal behavior "may be reinforced because it helps in the acquisition of other types of verbal operants" (p. 67). Facilitative reinforcement functions as a type of automatic reinforcement, like that observed in emitting a chain of verbal behaviors in problem solving. The successful acquisition of the next step in the chain is facilitated by an earlier step. Such facilitative reinforcement plays a causal role in more complex behavioral relations such as joint control and BiN. For example, repeating instructions (e.g., turn left at the 7-Eleven) while driving can facilitate the emergence of joint control when the 7-Eleven is encountered (Lowenkron, 1998).

Emergent Intraverbal Relations

Research has demonstrated that intraverbal relations can emerge from previously acquired tact, intraverbal, and listener repertoires, and even from direct observation (e.g., Allan, Vladescu, Kisamore, Reeve, & Sidener, 2014; Braam & Poling, 1983; Devine et al., 2016; Grannan & Rehfeldt, 2012; Pérez-González & García-Asenjo, 2016; Pérez-González, García-Asenjo, Williams, & Carnerero, 2007; Petursdottir, Carr, Lechago, & Almason, 2008; Smith et al., 2016). There are also ways that new intraverbal behavior can emerge from the combinations of existing verbal repertoires (DeSouza, Fisher, & Rodriguez, 2019; Horne & Lowe, 1996).

Intraverbal Bidirectional Naming (I-BiN)

Horne and Lowe (1996) suggested that emergent intraverbal relations could also be produced through intraverbal bidirectional naming (I-BiN) procedures (Miguel, 2016). In I-BiN a bidirectional relation is established between at least two verbal stimuli (e.g., "cat" and "animal"). I-BiN occurs when AB training on a unidirectional intraverbal relation automatically produces a reverse BA intraverbal relation, and vice versa. For example, after learning the unidirectional intraverbal relation between "cat" as the verbal S^D and "animal" as the verbal response, without further training a child can emit "cat" as the response when "animal" is the verbal S^D. The components of I-BiN consist of a combination of echoic, self-echoic, listener, self-listener, and tact repertoires working together. Once established, I-BiN, like C-BiN, can function as a higher-order generalized verbal operant that can produce powerful generative effects (e.g., Horne & Lowe, 1996; Greer & Ross, 2008; Jennings & Miguel, 2017; Miguel, 2016, 2018; Pérez-González et al., 2018).

Simple I-BiNs can be established by teaching a child a relation between two verbal stimuli (e.g., "shoes and socks"), and then teaching the child to emit overt or covert self-echoic repetitions of the trained intraverbal relation (e.g., "shoes and socks, shoes and socks"). When repeated, the basic components of the reverse intraverbal are contained contiguously in the phrase (i.e., the word "shoes" now follows the word

BOX 18.1

I Shouldn't Have Opened My Big Mouth: Skinner's Analysis of Self-Editing

Some people may say things that are offensive, socially inappropriate, or untruthful. Others may make frequent verbal mistakes, tangle their words, or digress repeatedly. Skinner devoted three chapters to self-editing (Chapters 15 to 17) in his book *Verbal Behavior* (1957). He emphasized that a speaker is also a listener, including a listener of her own verbal behavior and may, to her benefit, self-edit that verbal behavior before or soon after it is emitted. Self-editing is extra work for a speaker and occurs due to a history of reinforcement or punishment from listeners. However, Skinner notes that in many cases punishment may play a larger role in self-editing than reinforcement.

Some types of language and social problems can be related to self-editing deficits. Skinner (1957) suggests two behavioral distinctions be made when considering self-editing problems. The first is between a speaker who cares about what she says and a speaker who does not care about what she says. Caring involves a motivating operation (MO) that is related to a history of reinforcement and punishment from listeners. The same speaker may care that her verbal behavior obtains a positive listener response from one audience, but not from another audience. Caring can also involve MOs related to negative responses from listeners. A speaker may care that her words hurt others.

A second distinction can be made between a speaker who is aware of what she says and its effect on listeners and a speaker who is unaware of what she says or its effects. This distinction is based on the degree to which a speaker can tact aspects of her own verbal behavior (e.g., that certain words might be offensive) and aspects of her listener's behavior (e.g., noticing a nonattentive listener), and behave accordingly. Three examples of self-editing problems based on these distinctions are presented, along with some comments regarding intervention.

Offensive Verbal Behavior

Some individuals are aware of the offensive, hurtful, or socially negative things they say to people, but they may not care. They may demonstrate weak MOs for positive reactions from listeners, strong MOs for negative reactions from listeners, or strong MOs for social reinforcement from peers for their offensive behavior. People who say negative things to other people may not seek treatment because they do not feel anything is wrong. It may be hard to change this behavior because of the difficulty in controlling the relevant contingencies, especially the MOs.

Weak Verbal Repertoires

Some individuals are aware of their difficulties in expressing themselves. They might be insecure about what they say, tangle their words, or struggle with public speaking. Individuals experiencing these problems typically care about what they say, but they may have weak intraverbals, fail to tact listener behavior, or have interfering MOs and emotional responses. These individuals are more likely to seek treatment (e.g., Toastmasters, public speaking courses). Skinner (1957, pp. 405–417) describes several techniques for teaching a willing participant to edit her own verbal behavior (e.g., rehearsal with familiar audience, self-prompts, practice self-listening, practice tacting listener behavior, change the audience, change the motivation).

High Rate of Verbal Behavior

Some people may be unaware of the effects of their words on listeners, but they do care (a strong MO) that people respond positively. These individuals may engage in a high rate of talking, dominate a conversation, or provide excessive detail regarding events or their topic of interest. Typically, punishment reduces this behavior, but there may not be enough punishment, or other listeners intermittently reinforce the behavior. These speakers may be automatically reinforced by their own verbal behavior, have weak listener repertoires, or have weak MOs for the listener's point of view. Individuals experiencing these types of editing problems may not seek treatment because they are unaware that a problem exists, although intervention could be successful.

Summary

Skinner's three chapters on self-editing provide a behavioral analysis of many types of self-editing issues and problems. We may open our big mouths, then wish we had not, because we failed to edit our verbal behavior before it was emitted. Self-editing requires extra work for a speaker and may not occur for that reason. Skinner (1957) offers many practical suggestions for clinicians and others interested in understanding or improving self-editing repertoires.

"socks"), and this response is followed by direct or automatic reinforcement (Vaughan & Michael, 1982). When this reverse intraverbal emerges without training, this I-BiN effect is consistent with symmetry and mutual entailment (e.g., Horne & Lowe, 1996; Miguel & Petursdottir, 2009; Pérez-González et al., 2007; Petursdottir & Haflidadóttir, 2009). Once simple bidirectional relations are established between words, new words can be added to the class (e.g., "feet" "shirt") and ever more forms of emergent intraverbal behavior can occur (e.g., saying "feet" upon hearing "shoes go on your . . . ").

When a stimulus change affects one member of an I-BiN class, its effects can transfer to all members of the class.

For example, if a child has acquired the C-BiNs "dangerous" and "spider," and then learns the I-BiN "spiders are dangerous," this new verbal stimulus can produce a generative verbal function-altering effect (Schlinger & Blakely, 1994) for the whole stimulus class of spiders. For example, later when the child sees a spider, if she tacts the spider the response product of the tact (i.e., hearing the word "spider") evokes the intraverbal response "dangerous," bringing with it a history of listener behavior. Furthermore, the nonverbal stimulus of the spider is present when reinforcement is delivered, and as a result the tact of the spider as "dangerous" may also emerge. The new tact and intraverbal relations involving the word "dangerous" are likely members of several different equivalent and non-equivalent classes, and their effects can transfer throughout those classes (Chapter 20). For example, the previously neutral spider may now function as a conditioned stimulus and elicit respondent behaviors or as a CMO-R that evokes avoidance behaviors or as a nonverbal S^D that evokes the emergent tact "dangerous."

The verbal function-altering effect can also transfer to any stimulus that is in the same equivalence class as spiders (e.g., mosquitoes, ants, crickets) and similar respondent, avoidant, and tacting behavior can emerge. What makes this transformation of function important to generative language is that it demonstrates how an emergent tact relation can produce an emergent intraverbal relation, further demonstrating the far-reaching effects of combining verbal cusps (Greer & Ross, 2008; Hayes et al., 2001; Horne & Lowe, 1996; Rosales-Ruiz & Baer, 1997).

There is a growing body of research on more complex I-BiNs and their generative effects (e.g., Carp & Petursdottir, 2015; Greer & Ross, 2008; Jennings & Miguel, 2017; Ma, Miguel, & Jennings, 2016; Pérez-González et al., 2018; Santos, Ma, & Miguel, 2015). For example, Jennings and Miguel (2017) examined the effects of intraverbal and tact training on the emergence of stimulus equivalence. These researchers first taught 17 adult participants to tact birds (A), trees (B), and reptiles (C). Then the experimenters taught them unidirectional intraverbal relations between specific sets of stimuli (e.g., "The tree for Cardinal [A1] is Buckeye [B1] and the reptile for Buckeye [B1] is Black Racer [C1]"). Participants were then administered a visual-visual MTS test and an intraverbal test to assess performance for equivalence class formation. In addition, participants experienced four novel sets of stimuli to assess for generalization. The results for 4 of the 17 participants are presented in Figure 18.5. The closed circles, squares, and triangles represent three types of matching-to-sample tasks that constitute equivalence. Baseline data on equivalence show that these participants responded at chance level. However, following tact and intraverbal training, all 4 participants demonstrated nearly 100% correct performance on all equivalence measures. In addition, all the participants demonstrated emergent intraverbal behavior that participated in establishing the equivalence relations. In total, 13 of the 17 participants demonstrated the emergence of stimulus equivalence. These authors concluded that the participant's equivalence performance was established by the participant's own C-BiN and I-BiN verbal behavior.

LISTENER BEHAVIOR

Skinner's *Verbal Behavior* (1957) was primarily about the behavior of the speaker, but Skinner argued that a complete account of language must include both speaker and listener behavior. In contrast, linguists tended to ignore the speaker and were mainly concerned with the listener (Skinner, 1978). Gradually however, Skinner's book began to affect linguistic theory (Andresen, 1991). In a letter Skinner sent to me in 1989 (see Box 18.2) he muses, "It is quite amazing how linguists are coming around to the position of my book. They had to get around to the behavior of the speaker sooner or later, having spent centuries on how verbal behavior is understood by the listener." Despite Skinner's focus on the speaker, Schlinger (2008b) noted the following: "Skinner did not neglect the listener. In fact, he made frequent mention of the listener throughout the book—the word *listener* occurs 793 times compared to 893 instances of the word *speaker*" (p. 312). In addition two chapters of *Verbal Behavior* are primarily devoted to analyses of listener behavior (Chapters 6 and 7).

Different Roles of a Listener

A **listener** is an active participant in verbal exchanges and serves many roles in this capacity. Skinner (1957) makes a distinction between the nonverbal and verbal duties of a listener. The nonverbal duties include serving as an audience (an S^D or MO) and providing reinforcement for speakers. Reinforcement for a speaker typically involves some kind of listener action directly controlled by the speaker's words. Listener action can be nonverbal (e.g., a child gets his shoes when asked to do so) or it can be verbal (e.g., a child says "Can you help me?"). When a listener takes verbal action, she now becomes a speaker, and her behavior should be interpreted as such (e.g., a mand). Speakers and listeners reverse roles, often rapidly in verbal exchanges, and usually speakers and listeners are also self-listeners of their own verbal behavior. In any given verbal exchange, all of these repertoires interact with each other and with the current context and contingencies. The interaction between speakers and listeners (including self-listeners) constitutes what Skinner termed a **verbal episode**, and can be considered the basic speaker-listener unit of verbal behavior. The roles of the listener and the elements of a verbal episode will be described in more detail.

Mediator of Reinforcement

One role of a listener is to consequate a speaker's behavior, which can occur in a number of ways. Eye contact, head nods, social praise, smiles, and so on can all function as generalized S^r. But listeners do more than just directly reinforce a speaker for emitting verbal behavior; they reinforce the speaker's behavior by acting on her verbal behavior. For example, if a speaker says, "Can you hand me the remote?" the listener's nonverbal behavior of passing over the remote functions as specific reinforcement for the speaker's verbal behavior and abolishes the MO. A listener may also ignore or punish a speaker's verbal

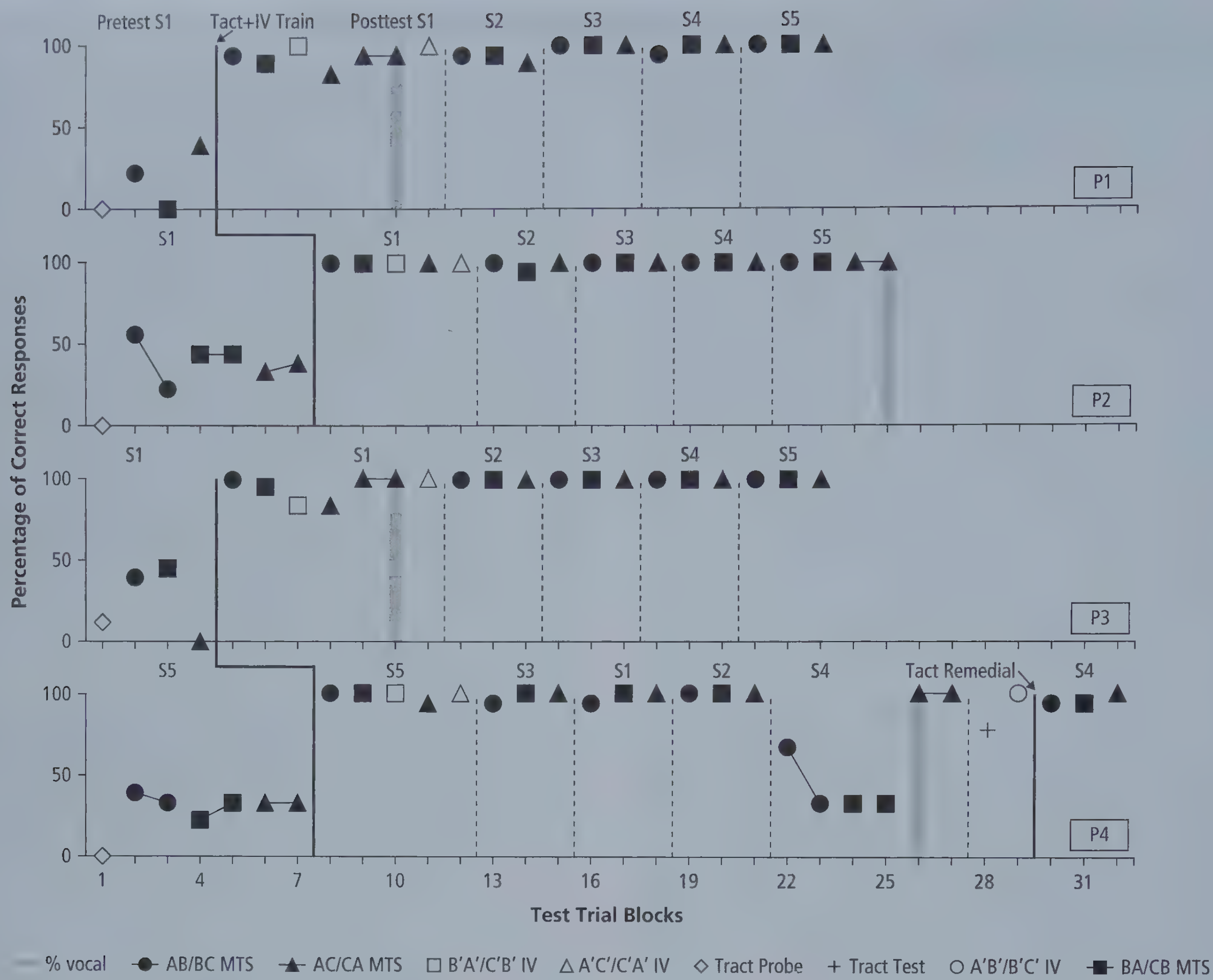


Figure 18.5 Percentage of correct responses during matching-to-sample (MTS) tasks, tacts, listener, and intraverbal tests across five sets of stimuli for 4 of the 17 participants. The closed circles, squares, and triangles represent three types of MTS tasks that constitute equivalence as measured in baseline conditions, followed by performance after tact and intraverbal training, and then across five generalization sets. The open circles and squares represent measures of emergent bidirectional intraverbal behavior.

From "Training Intraverbal Bidirectional Naming to Establish Generalized Equivalence Performance" by A. M. Jennings and C. F. Miguel, *The Journal of the Experimental Analysis of Behavior, 108,* p. 279. Reproduced with permission of John Wiley & Sons, Inc.

behavior, likely resulting in an extinction burst or decrease in responses. These types of consequences delivered by listeners shape and maintain speaking behavior, and help bring it under S^D and MO control.

Serving as an Audience

The listener also serves as a nonverbal S^D for a speaker's behavior. Skinner (1957) devotes an entire chapter to the role of the listener as an audience, and in doing so identifies three subdivisions of an audience. The first and largest subdivision is the language spoken (e.g., an English-speaking audience evokes English words). The second level involves special audiences who evoke differential responses. For example, one might use the term "reward" when talking with parents and laypersons but the term "reinforcement" when talking with behavior analysts. The third and most extensive role of the audience is to "select a subject matter" (Skinner, 1957, p. 175). We talk about different topics with different people. For example, a discussion involving sports may be reinforced by one member of a speaker's verbal community, but punished by another, and each audience acquires multiple and generative discriminative functions.

Understanding a Speaker

The most complex aspect of a listener's behavior involves what can be identified as understanding what a speaker says. Understanding on the part of a listener can be demonstrated by nonverbal or verbal behaviors. Nonverbal understanding can be

B. F. Skinner's Letter To Mark Sundberg

BOX 18.2

HARVARD UNIVERSITY

DEPARTMENT OF PSYCHOLOGY

WILLIAM JAMES HALL
33 KIRKLAND STREET
CAMBRIDGE, MASSACHUSETTS 02138

August 7, 1989

Dr. Mark L. Sundberg

Dear Mark:

Thanks for the copy of the new edition. I thought there were some excellent things in it. I like the review of Bruner. It is quite amazing how the linguists are coming around to the position of my book. They had to get around to the behavior of the speaker sooner or later, having spent centuries on how verbal behavior is understood by the listener.

The book you cited must have been my <u>Recent Issues in the Analysis of Behavior</u>, not anything to do especially with verbal behavior.

Things are going well here. I am transcribing and editing a draft of my paper at ABA for the <u>Journal of Behaviorology</u>.

All the best,

B.F Skinner

B. F. Skinner's Letter To Mark Sundberg HARVARD UNIVERSITY DEPARTMENT OF PSYCHOLOGY. Used with permission

demonstrated by a verbal stimulus that evokes a corresponding nonverbal response. For example, in standard receptive language training, a child may be shown an array of buildings and asked, "Where is the post office?" The nonverbal selection of the picture of a post office would demonstrate a type of understanding. Verbal understanding can be demonstrated by a verbal stimulus that evokes an appropriate verbal response from the listener (who is now a speaker). For example, if a tourist asks a passerby for the location of a post office, and the passerby says, "About three blocks down, on the left," understanding what the tourist asked is demonstrated by the intraverbal behavior emitted by the passerby.

The same four types of verbal discriminations previously described for intraverbal relations can be used as a framework to identify various types of nonverbal listener discriminations (Figure 18.4). In addition, there is a fifth type of verbal discrimination that is specific to listener behavior commonly termed an auditory conditional discrimination (e.g., Saunders & Green, 1999; Sidman et al., 1982). Each of these types of listener discriminations demonstrates how multiple control in the form of a conditional discrimination between verbal and nonverbal stimuli is an important aspect of listener understanding.

Speaker and Listener Verbal Episodes

Figure 18.6 presents an example of a verbal episode in which a speaker in a restaurant asks a waiter, "May I have some bread?" The sources of antecedent control are multiple: (1) an MO related to bread (MO_1), (2) the general context (the restaurant) (nonverbal S^{D}_{1}), (3) a receptive audience (the waiter) (nonverbal S^{D}_{2}); if the waiter asks, "Do you need anything else?" these (4) verbal stimuli are relevant (verbal S^{D}_{1}), and if bread is visible, that (5) nonverbal stimulus (nonverbal S^{D}_{3}) could also participate in the antecedent configuration. For the listener (the waiter), his behavior is controlled by the speaker's words, "May I have some bread?" (verbal S^{D}_{2}). The speaker's words can have multiple divergent effects on listener behavior. They can evoke a verbal response (e.g., "Yes, of course") and a nonverbal response (e.g., the waiter delivering bread), but only the delivery of bread functions as the ultimate specific reinforcement for the speaker, and changes the function of the MO to an abolishing operation (AO). Generalized S^r in this example, although encouraging, would not satisfy the MO. There can be other variables involved in this episode as well (e.g., listener MOs, respondent relations, covert behaviors).

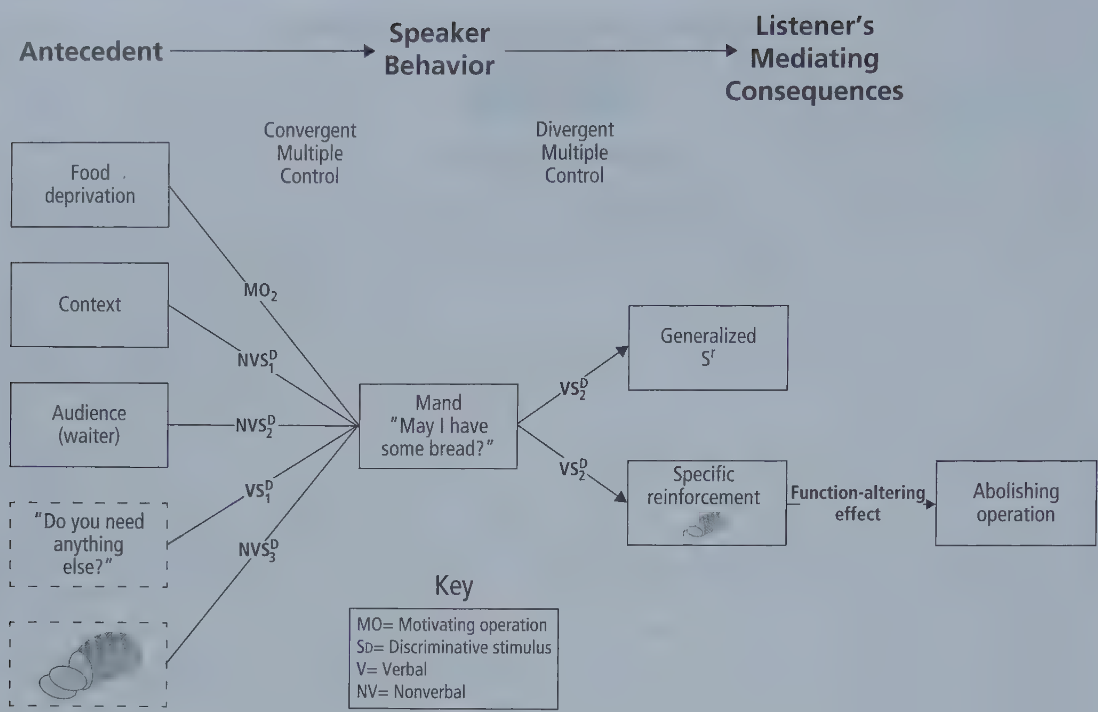


Figure 18.6 A verbal episode between a speaker and a listener.

This interaction could also generate new MOs and S^Ds (e.g., butter, a glass of wine).

Figure 18.6 also demonstrates how the same word can have different meanings for speakers and listeners, and why separate accounts are necessary. In this example, the same word "bread" and its nonverbal referent are involved for each party, but the word "bread" appears in different positions in the individual four-term contingencies affecting the speaker and listener. For the speaker, "bread" is a response, but for the listener, "bread" is a stimulus. In addition, the word has different functions for each participant. For the speaker, "bread" is a response in a mand relation, but for the listener "bread" is a verbal S^D that evokes a nonverbal discrimination. Thus, what the word "bread" means for the speaker is not the same as what the word "bread" means for the listener, despite having the same formal properties and nonverbal linguistic referent.

Emergent Listener Relations

A listener is an active responder, and new listener relations can emerge through the process of stimulus and response generalization. In addition, new relations can emerge through generative processes produced by the combinations of previously acquired skills. Several studies have demonstrated that novel listener behaviors can emerge from direct training on speaker relations (e.g., Byrne, Rehfeldt, & Aguirre, 2014; Cuvo & Riva, 1980; Ingvarsson, Cammilleri, & Macias, 2012; Lowe, Horne, Harris, & Randle, 2002; Petursdottir & Carr, 2011; Petursdottir & Haflidadóttir, 2009). For example, Ingvarsson et al. (2012) showed that direct intraverbal training was sufficient to generate emergent listener discriminations for children with autism.

Common Bidirectional Naming (C-BiN)

The merging of speaker and listener cusps is the foundation of BiN. Once a child has acquired the basic echoic, tact, listener, and self-listener cusps, the prerequisite repertoires for C-BiN are established and a novel word acquired (or observed) as a tact can immediately generate a listener relation without further training (i.e., symmetry, mutual entailment). Several studies have demonstrated this type of bidirectional naming and its effects on emergent listener behavior (e.g., Gilic & Greer, 2011; Greer & Ross, 2008; Horne & Lowe, 1996; Miguel & Kobari-Wright, 2013).

Conditioning the Behavior of the Listener

Emergent listener behavior can also occur through the behavior of a speaker who provides instructions (rules) that condition the immediate or future behavior of a listener (Schlinger, 2008b; Skinner, 1957, 1969). A speaker's words can alter the functional properties of other words, objects, and events for a listener, but not until those stimuli are encountered (Schlinger & Blakely, 1994). For example, if a speaker says, "Don't go to Bill's Restaurant—it's horrible," any stimulus in that equivalence relation can acquire an emergent evocative effect if encountered. Bill's Restaurant may now evoke avoidant behavior or the tact "horrible." In addition, the restaurant, its owner, or patrons may now evoke behavior in equivalence relations with horrible things, which might include aggression or property destruction (e.g., vandalism). These transformation of function effects demonstrate the generative effect of instructions, and show the potential impact that words can have on behavior.

AUTOCLITIC VERBAL BEHAVIOR

In the process of acquiring language, a child learns to listen to the words of others. Along the way, the child also learns to listen and react to her own words. She may self-edit, qualify, order, negate, or elaborate on her words. This is an important and complex secondary type of verbal behavior that is shaped and reinforced by listeners in a speaker's verbal community. For example, a supervisor may pause briefly before saying to a staff member, "That was good." What might be the behavioral effects of the pause? Does it have any special meaning to a listener? The pause can produce a verbal stimulus that accompanies what the listener hears, and it can change the way a listener reacts to what he hears. Speakers manipulate their own verbal behavior, often accidentally or unknowingly, because such behavior is reinforced by listeners. Skinner terms this secondary verbal behavior autoclitic behavior.

The **autoclitic** relation involves "two systems of responses, one based upon the other. The upper level can only be understood in terms of its relation to the lower level" (Skinner, 1957, p. 313). The upper level may be described as a speaker's verbal behavior about her own lower level verbal behavior. Michael (1991, 1992) suggested that these two levels of responses be identified as *primary verbal behavior* as the lower level and *secondary autoclitic behavior* as the upper level. A primary verbal response constitutes any of the verbal operants (e.g., mand, tact, intraverbal, textual), and a secondary autoclitic response is controlled by "some feature of the primary verbal operant or by its controlling variables" (Michael, 1992, p. viii). For example, a primary response such as "The ice is solid" may be accompanied by a secondary response, "I think," that informs the listener of the tenuous nature of the primary tact. Autoclitic behavior is helpful to listeners who may prompt a speaker to identify the sources of control for what she says, as in "Did you test the ice?" The autoclitic response tells the listener more than just the primary response on its own, and listeners reinforce speakers for this extra information.

Michael (1991) suggests that not all of a speaker's secondary verbal behavior has an autoclitic function. Some of a speaker's response to her own verbal behavior consists simply of more primary verbal behavior. For example, covert verbal behavior regarding an upcoming meeting may evoke self-intraverbals or self-mands regarding that meeting. Michael (1991) identifies this type of response to one's own verbal behavior as *simple secondary verbal behavior*. There are several ways to distinguish between secondary autoclitic responses and simple secondary responses. For example, a secondary autoclitic response cannot affect behavior on its own; it must involve some aspect of a primary response. Simple secondary behavior, requires a strict temporal relation between the primary response and the secondary response, with the primary response always occurring first and generating the relevant verbal S^Ds that evoke the secondary verbal behavior. The autoclitic relation differs in that the primary response and secondary autoclitic response can occur in any order, or simultaneously (e.g., "I think that's the right exit," "That's the right exit, I think," or slowly emitting "right"). Peterson and Ledoux (2014) suggested that secondary autoclitic relations be further classified as autoclitic mands and autoclitic tacts.

Autoclitic Mand

The **autoclitic mand** is a type of secondary autoclitic behavior that has the same defining features as a primary mand (responses controlled by an MO and specific reinforcement), except the autoclitic mand is secondary verbal behavior that enjoins a listener to take some specific action regarding the primary response, and autoclitic behavior is reinforced by the listener's action. For example, a speaker who says, "Believe me, they are wrong," is blatantly manding to the listener with "believe me" to verbally accept as fact, without challenge, her statement "They are wrong." The autoclitic mand can take any form, accompany any type of verbal interaction, and range from subtle to obvious.

Autoclitic Tact

An **autoclitic** tact is a type of secondary verbal behavior that is controlled by some nonverbal feature of the primary response or its controlling variables, and the autoclitic response informs the listener of that feature. This behavior is reinforced by listeners who provide generalized S^r. For example, a speaker may say, "I'm sure it's safe," where the secondary autoclitic response "I'm sure " is a tact of the strong source of control for "it's safe." Skinner (1957) identified several types of autoclitic tacts. A *descriptive autoclitic tact* is controlled by the nonverbal properties of the antecedent events and their relation to the primary response. There are several types of descriptive autoclitics. For example, an autoclitic tact of strength could be observed in a speaker who might have a tendency to say, "That's the right house" (a primary response), and covertly tacts the strength of the nonverbal stimulus control (a secondary autoclitic response) and tags on, "I'm sure," before "That's the right house." "I'm sure" is a descriptive autoclitic tact that informs the listener of the strength of the primary relation and as a result alters the listener response to "That's the right house." The listener reinforces the speaker's behavior because he is benefited by this additional information.

Quantifying autoclitic tacts inform the listener of the range of application of a primary response. For example, a speaker may tag on the autoclitic tact "some" to the primary tact "Behavior problems are caused by attention." The autoclitic "some" informs the listener that the whole statement that follows is restricted in its range. If "all" had been tagged on to the primary response, the listener would behave differently. *Relational autoclitic tacts* inform the listener of agreement between aspects of the primary responses. For example, possessive relations indicated by an *'s* can be tagged to a noun, as in the "supervisor's recommendation," and function as a tact of the source of control for the recommendation, and as a tact of the agreement between the source and the recommendation. Skinner suggested that autoclitic relations play a causal role in complex verbal activities involving grammar, syntax, composition, grouping, ordering, punctuation, predication, and so on (e.g., Palmer, 2016).

Intraverbal Autoclitic Frames

An intraverbal autoclitic frame facilitates the relation between words in a given utterance by providing order, agreement, grouping, and composition of larger units of verbal behavior. Moore (2008) describes an autoclitic frame as "conventionally prescribed formats of sequences for emitting verbal behavior, as in subject-verb-object or actor-action-object sentences. These frames may become very generalized, such that they come to be applied in many other situations than the original" (p. 208). Skinner (1957) provides the following example of an autoclitic frame:

> If (a boy) has acquired a series of responses such as *the boy's gun, the boy's shoe*, and *the boy's hat*, we may suppose that the partial frame *the boy's* _____ is available for recombination with other responses. The first time the boy acquires a bicycle, the speaker can compose a new unit *the boy's bicycle*. . . . The relational aspects of the situation strengthen a frame, and specific features of the situation strengthen the responses fitted into it" (p. 336).

Palmer (2016) notes "many of the structural properties of verbal behavior that so excite the linguist arise, at least in part, from the prevalence of intraverbals and intraverbal control. In particular, autoclitic frames and grammatical tags are largely intraverbal" (p. 101). Speakers must acquire these intraverbal autoclitic frames and tags because they play an important role in the effective execution of verbal behavior. Very little research on autoclitic frames is available, but some has begun to emerge (e.g., Luke, Greer, Singer-Dudek, & Keohane, 2011; Martins, Hübner, Gomes, Pinto Portugal, & Treu, 2015; Speckman, Greer, & Rivera-Valdes, 2012).

APPLICATIONS OF SKINNER'S (1957) ANALYSIS OF VERBAL BEHAVIOR

One of the most developed applications of Skinner's work can be found in the behavioral treatment of autism or other types of intellectual disabilities.[4] Linguistic problems are often the primary handicapping conditions of those diagnoses, and effective language assessment and intervention strategies are essential for treatment programs. Joe Spradlin was the first to publish an application of the concepts from *Verbal Behavior* to the communication needs of the intellectually disabled. He created the *Parsons Language Sample* (1963), which was an assessment tool based on the verbal operants. Spradlin was also instrumental in the early research and development of language intervention programs based on verbal behavior (e.g., Spradlin, 1966). Jack Michael was one of the first professors to offer a full course using Skinner's (1957) book *Verbal Behavior*, and he spearheaded the further development of Skinner's analysis of verbal behavior (see Box 18.3).

BOX 18.3

Jack Michael and Skinner's Analysis of Verbal Behavior

Jack Michael began teaching behavior analysis in 1955 at the University of Kansas. He used Skinner's (1953) *Science and Human Behavior* as the textbook for the course. In that book, Skinner mentioned his upcoming book *Verbal Behavior* in a footnote (p. 210). Jack contacted Skinner regarding the book, and Skinner sent him an early version of *Verbal Behavior*. Jack began to incorporate verbal behavior content into his courses. After the book became available, Jack developed a full course on verbal behavior and taught that course nearly every academic year until his retirement in 2003. He maintained, like Skinner (1953), that an analysis of verbal behavior was essential for a complete account of complex human behavior (e.g., Michael, 1984).

One of Jack's primary goals as a teacher was to impart to students the verbal repertoires necessary to precisely use the concepts and principles of behavior analysis to analyze behavior, including verbal behavior, in any context (Michael, 1995, 2004). In the process of teaching others, he was constantly advancing his own understanding of Skinner's writings while refining various concepts along the way (e.g., motivating operations, positive and negative reinforcement, automatic reinforcement, topography-based and selection-based verbal behavior, duplic and codic relations).

Applied research on verbal behavior was important to Jack, but it was slow in developing, despite Skinner's (1957) prediction that "[t]he formulation is inherently practical and suggests immediate technological applications at almost every step" (p. 12). In 1976, Jack started offering a graduate course at Western Michigan University titled "Verbal Behavior Applications." That course inspired many students, especially those needing theses or dissertations, to conduct verbal behavior research. The Midwestern Association for Behavior Analysis (MABA) convention provided contingencies to present the research, and a verbal community arose that eagerly supported verbal behavior research.

Jack and W. Scott Wood were instrumental in starting the Verbal Behavior Special Interest Group (VB SIG). They co-chaired the first meeting, which was held at the 1977 MABA convention. Skinner attended and participated in the discussions, as did many other prominent behavior analysts. It was suggested that an outlet for verbal behavior research and other instructional material was needed. Under Jack's direction, the VB SIG established a newsletter, the *VB-NEWS*, which ultimately became the journal *The Analysis of Verbal Behavior*. Thanks to Jack's relentless dedication to understanding, teaching, advancing, and applying Skinner's analysis of verbal behavior, we now have a robust body of conceptual and empirical research, and many highly successful clinical applications.

APPLICATIONS TO LANGUAGE ASSESSMENT AND INTERVENTION

Behavior analysis provides three major contributions to language assessment and intervention programs for individuals with special needs: (1) applied behavior analysis technology, (2) a behavioral analysis of human development, and (3) a behavioral analysis of language.

Applied Behavior Analysis (ABA) Technology

Behavioral technology provides teachers, speech-language pathologists (SLPs), parents, and others with the necessary procedures and strategies to assess and teach language skills to those who need formal instruction (e.g., Leaf & McEachin, 1999; Lovaas, 1987, 2003; Maurice, Green, & Luce, 1998; Romanczyk & McEachin, 2016). Effective use of behavioral procedures requires specific training and supervision. For example, staff must know when and how to use reinforcement, extinction, prompting, fading, and so on. The content of the current book describes the essential features of ABA, as do a number of other behavior analysis textbooks (e.g., Malott & Shane, 2013; Martin & Pear, 2015; Miltenberger, 2016; Sulzer-Azaroff, Mayer, & Wallace, 2013).

A Behavioral Analysis of Human Development

The development of language and social skills of typically developing children provides a valuable guide and framework for working with children with language delays. Knowing what skills to teach and when those skills appear in typical language development is an important component of establishing and sequencing a language intervention program. Behavior analysis has a long history of child development research, beginning with the books by Bijou and Baer (1961, 1965, 1967). Their work provides a behavioral analysis of how children learn and develop skills throughout their early childhood. There is a dedicated journal (*Behavioral Development Bulletin*) and numerous extensions and applications of a behavioral approach to human development (e.g., Greer & Koehane, 2005; Morris et al., 1982; Novak & Pelaez, 2004; Rosales-Ruiz & Baer, 1997; Schlinger, 1995).

A Behavioral Analysis of Language

Skinner's (1957) analysis of verbal behavior provides a framework for language assessment and intervention. A verbal behavior assessment program involves identifying the strength of the elementary verbal operants, listener skills, self-listener skills, multiply controlled responding, and speaker-listener interactions in a developmentally sequenced fashion. Intervention is based on the assessment results and focuses on establishing missing or impaired speaker or listener skills through the use of ABA teaching procedures. Several publications are now available that provide detail on how to apply Skinner's analysis of verbal behavior in working with individuals who have language delays (e.g., Barbera & Rasmussen 2007; Carr & Miguel 2012; Dixon, 2014; Greer & Ross, 2008; LeBlanc & Dillon, 2009; McGreevy et al., 2012; Partington, 2006; Schramm, 2011; Sundberg, 2014, 2016b; Sundberg & Michael, 2001; Sundberg & Partington, 1998; Weiss & Demiri, 2011).

The Importance of Language Assessment

Parents and pediatricians usually monitor a child's acquisition of language, noting important milestones while watching for discrepancies. Language delays are associated with a variety of diagnoses (e.g., autism, communication disorders) and early identification is valuable. If a child cannot communicate with words, negative behavior (e.g., tantrums, aggression, social withdrawal) may serve the communication function. When a child's language skills fail to develop, or are impaired in some way, professional assistance should be obtained, such as services from an SLP, behavior analyst, or psychologist.

A language assessment for a person with language delays has three primary goals: (1) to identify the nature of the delay or problem, (2) to compare the person's performance with expected norms, and (3) to provide guidance for an appropriate intervention program if necessary. Each type of speaker and listener behavior needs to be assessed independently and in various combinations (e.g., C-BiNs, equivalent and non-equivalent relations) to obtain baseline levels of performance and to identify the direction of an intervention program. The goal is to determine what a child can and cannot do (e.g., mand, echo, listen, tact, demonstrate generative learning). The value of assessing each of the speaker and listener domains, and the interaction between them, is that it is common to encounter language-delayed children who can emit a word as one type of verbal behavior (e.g., a tact) but not as another (e.g., a mand).

A variety of standardized language assessments are available, but none provide measures of all of the verbal operants and listener skills (Esch, LaLonde, & Esch, 2010). However, several criterion-referenced language assessments are available that do measure each of the verbal operants and listener repertoires (e.g., Dixon, 2014; McGreevy et al., 2012; Partington, 2006; Partington & Sundberg, 1998; Spradlin, 1963; Sundberg, 1983, 2014; Sundberg & Partington, 1998; Sundberg et al., 1979). These assessments are framed by the elementary verbal operants, listener behavior, various combinations of speaker and listener behavior, and perceptual behavior. In the following sections, early language assessment and intervention strategies will be described for each of the individual speaker and listener domains.

Mand Assessment

Assessing a child's mand repertoire involves identifying the degree to which MOs control verbal responses. The mand classification system previously described can guide the assessment process. Specifically, the distinctions between the types of MOs (i.e., UMO, CMO-T, CMO-R) and their evocative effects on verbal behavior can be assessed. For example, does the child emit a mand when deprivation UMOs are at strength (e.g., saying "milk" when she wants milk) or when aversive stimuli function as UMOs (e.g., when it is cold, saying "Let's go in"). A similar process is necessary for learned MOs such as the CMO-T (e.g., manding for a missing toy) and the CMO-R (e.g., manding to fix

a broken toy). In addition, the delivery of specific reinforcement should abolish (AO) the CMO-R (e.g., the toy is fixed) and this effect should be assessed.

Mands can also be assessed across MOs relevant to the different parts of speech (e.g., MOs related to action, location, possession) and in social interactions between speakers and listeners (e.g., manding to peers). Also, Skinner's nine types of mands (Figure 18.3) can provide a further quantification of the different MOs that control mands for a specific person. Mands involving multiple variables should also be assessed, such as those involving multiple MOs (e.g., manding "I want scissors and Play-Doh"), MOs accompanied by nonverbal S^Ds (e.g., "I want that red one"), MOs accompanied by verbal S^Ds (e.g., manding for a specific video game after a salesperson's pitch), and MOs accompanied by both verbal and nonverbal S^Ds (e.g., asking a waiter for bread).

Barriers to Manding

An important component of an intervention program is an analysis of the language, learning, and social barriers demonstrated by a child during verbal interactions with others (Sundberg, 2014). The mand is more susceptible to interference from barriers than are the other verbal operants because of the mand's relation to MOs and specific reinforcement. These variables are powerful and can quickly evoke and shape negative behavior. For example, a child's scream may immediately evoke eye contact and attention from an adult, and as a result screaming behavior increases and comes under MO and S^D control. These naturally occurring contingencies may be more efficient than the carefully nurtured and programmed contingencies offered in a structured intervention program. There are several common barriers involving MOs and specific reinforcement, such as demand weakening the MO, quick satiation, competing MOs and S^Ds, response forms inconsistent with the MO, and generalized CMO-Rs dominating behavior.

Mand Intervention

If a toddler fails to develop a mand repertoire, specific intervention may be necessary. Without words, signs, or icons as mands, the probability of negative behavior (e.g., tantrums) functioning as mands increases. The initial focus of an intervention program for nonverbal learners is usually on both establishing acceptable mand forms (preferably speech, but signs and icon selection can also be effective) and reducing existing forms of negative behaviors as mands.

Response Form

An important aspect of mand training for a nonverbal child is to establish a reliable response form that can produce the best possible effects on a listener. In early mand training there are three possible choices: speech, sign language, and icon-selection systems (in some special cases, written words can also be used as a fourth option). The primary focus should be on speech as a response form. However, if a child fails to acquire speech, sign language and icon-selection systems can work as mand forms (e.g., Frost & Bondy, 1994; Sundberg & Partington, 1998). Determining which form is most appropriate for an individual child depends on a variety of factors that should be carefully considered (Bondy, 2012; Carbone, Sweeney-Kervin, Attanasio, & Kasper, 2010; Sundberg, 1993a; C. T. Sundberg & Sundberg, 1990; Tincani, 2004).

Using MOs for Instructional Purposes

Teaching children with absent or impaired mand repertoires requires special skills on the part of caregivers. In addition to basic ABA knowledge, adults must be able to identify and control the different MOs and related specific reinforcement as part of their teaching skills. If the relevant MO is not at strength, early mand training is impossible. MOs are often difficult to control because they can vary in strength across time, be fleeting, have an instant or gradual onset or offset, and combine or compete with other MOs and S^Ds. In addition, too high a response requirement placed on a child may quickly weaken an MO (Alling & Poling, 1995).

Two general methods are used for identifying and controlling the value change of MOs for mand training: (1) capturing the MO value change as it occurs naturally and (2) creating the MO by generating some change that alters the value of a consequence (Michael, 1993; Sundberg, 1993b, 2013). Capturing motivation in the natural environment can be accomplished by observation of collateral behaviors that may be evoked by an MO (e.g., looking at a certain toy). Collateral behaviors provide insight as to the presence of a private MO, and thus an opportunity to teach a new mand to a child (e.g., "ball"). Creating motivation requires an adult to make some direct change (e.g., presenting a desired iPad without turning it on) that increases the value of some other change (e.g., turning on the iPad). The onset of this MO value change can be used to establish a new mand relation (e.g., "iPad on").

Capturing and Creating UMOs

The deprivation-based UMOs (e.g., food, drinks) are often an easier place to start for early or difficult learners because of the powerful evocative effects of these MOs (e.g., teaching "cracker" or "juice" as an early mand). UMOs can be captured for mand training purposes. For example, if a child reaches for a cracker, capture that MO and conduct a mand teaching trial. UMOs can also be created. For example, give the child some salty foods to increase the value of drinks as reinforcement. These conditions are also optimal for the implementation of pairing procedures to facilitate the development of additional behaviors, S^Ds, and forms of S^r (e.g., Esch, Carr, & Grow, 2009; Rader et al., 2014; Stock et al., 2008).

Capturing and Creating CMO-Ts

Capturing CMO-Ts in a child's natural environment can be accomplished by observing the events that increase the value of some object or event. For example, after an adult blows bubbles with a wand, a child demonstrates excitement while watching the bubbles float and pop. After they pop, the value of the wand and specific adult behavior increases, and that is the time to conduct a mand trial. Additional ways to capture MOs include the observation of established chains of behavior that are interrupted (e.g., missing toy piece), routines that are difficult (putting on

a jacket), and preferences (e.g., a certain cup). The pioneering work of Hart and Risley (1975) on incidental teaching exemplifies these teaching strategies.

CMO-Ts can be created by caregivers for language instruction as well (e.g., Albert, Carbone, Murray, Hagerty, & Sweeney-Kerwin, 2012; Hall & Sundberg, 1987; Landa, Hansen, & Shillingsburg, 2017; Oleson & Baker, 2014). For example, if a child likes to cut Play-Doh with scissors, then the presentation of Play-Doh may increase the value of scissors, demonstrated by the child's collateral behavior of reaching for scissors. The adult could block access to the scissors to control the MO and teach the mand. Establishing these mands can be accomplished by the use of prompting, fading, differential reinforcement, and transfer procedures (Sundberg & Partington, 1998).

Parts of Speech. The parts of speech provide a developmental framework for more advanced mand training. For example, in addition to mands for edible and tangible reinforcement, a child needs to learn mands for actions (verbs), for properties of objects (adjectives), for relations between objects (prepositions), and so on. All of these mands are controlled by different MOs, and each particular MO should be at strength at the moment of teaching.

Teaching Additional Types of Mands

Further direction for mand training can come from Skinner's list of nine types of mands (Figure 18.3). For example, a question is a mand controlled by an MO related to verbal information and specifically reinforced by the receipt of that information. Mands for information contribute to the burst in vocabulary growth commonly observed in toddlers between the ages of 2 and 3. These mands involve an increase in the value of the names of things and events (e.g., "What's that?") or the location of things and events (e.g., "Where's Buzz?"), while later questions may involve MOs for the function of things and events (e.g., "How do you turn it on?") or the cause of events (e.g., "How did it break?"). A growing body of research has provided support for the role of MOs in question asking and has demonstrated a variety of effective teaching procedures (e.g., Endicott & Higbee, 2006; Lechago, Carr, Grow, Love, & Almason, 2010; Lechago & Low, 2015; Shillingsburg, Gayman, & Walton, 2016; Sundberg, Loeb, Hale, & Eigenheer, 2002).

Multiple Control

MOs frequently share control with other MOs, verbal S^Ds, and nonverbal S^Ds as types of convergent multiple control. Multiple sources of control can be quite useful for establishing mands. A child's early manding may be controlled by MOs along with echoic, intraverbal, or nonverbal prompts, and through the process of differential reinforcement for higher quality responding, the response can be freed from these additional sources of control.

Teaching Extended Mands

Skinner (1957) identified several ways that mands generalize across MOs and S^Ds, and how the same MO can control different response forms (Miguel, 2017). Recent research has demonstrated methods to establish new mand forms for the same MO by using procedures from relational frame theory (e.g., Hayes et al., 2001; Murphy & Barnes-Holmes, 2009; Rosales & Rehfeldt, 2007). For example, Murphy & Barnes-Holmes (2009) established multiple types of derived mands involving "more" and "less" for three participants with autism. The study was conducted in the context of a board game involving the quantity of tokens required to be successful (more or less). These authors showed that a new mand response form could be established with a supplemental conditional discrimination training procedure that allowed for novel stimuli to become equivalent with existing mand forms. That is, the participants learned, without formal training, a new way to mand for "more" or "less."

Tact Assessment

Assessing a tact repertoire involves determining the degree to which nonverbal stimuli control verbal behavior. This is a sizable task given that the whole of the physical environment contains a massive number of nonverbal stimuli. The different parts of speech provide a frame for tact assessment because each part often demonstrates a different type of nonverbal stimulus control. Many nouns involve nonverbal S^Ds consisting of static objects, and many verbs involve nonverbal S^Ds consisting of movement, while many adjectives and adverbs involve nonverbal S^Ds generated by the properties of objects and events, and relations between them. A similar analysis can be made for the other parts of speech (e.g., prepositions, pronouns). In addition, an assessment of C-BiN relations can identify the strength of this higher order verbal operant (e.g., Greer & Ross, 2008; Miguel, 2016).

Sensory Receptors

Tacts can also be assessed across the different sensory receptors affected by nonverbal stimuli (e.g., visual, auditory, tactile)—for example, testing the tact of a cat when the target is a visual stimulus (e.g., a cat) versus an auditory stimulus (e.g., hearing a cat meow).

Multiple Control

Tacting becomes more complex when multiple variables are involved. For example, a noun-verb tact such as "The water's spilling out" involves at least two sources of stimulus control: One is a nonverbal S^D related to the water, and the other is a nonverbal S^D related to the movement of the water. Important information regarding a tact repertoire can be obtained by sampling the arrangement of various nonverbal stimuli corresponding to the different parts of speech (e.g., possessive-adjective-noun, subject-verb-noun). In addition, nonverbal S^Ds can combine with MOs and verbal S^Ds as types of convergent multiple control. For example, a child may emit the name of a video game she sees due to both an MO and a nonverbal S^D, making the response part mand and part tact.

An important aspect of tact assessment involves identifying the degree to which nonverbal stimulus generalization occurs. The four types of tact extensions provide a framework for assessing a person's skill in this area. In addition, the ability

to tact private events is a component part of a tact repertoire and can be measured by the occurrence of collateral behavior (e.g., limping) or public accompaniment (e.g., a bruise).

Tact Intervention

Learning to tact features of the physical environment is a cornerstone of early language acquisition. Once a tact repertoire is acquired, it can benefit a child in many ways. For example, tacting can facilitate listener behavior and the acquisition of other forms of verbal behavior such as mands, intraverbals, C-BiNs, equivalence classes, and relational frames (Horne & Lowe, 1996; Jennings & Miguel, 2017; Petursdottir & Carr, 2011; Sundberg, 2015). If a tact repertoire does not begin to develop for a toddler, intervention may be necessary. Teaching a child to tact is more straightforward than teaching a child to mand, because the antecedents that control the tact are more observable, measurable, quantifiable, and controllable than MOs. In addition, tacts are consequated by easily manageable and portable generalized S^rs, whereas the specific reinforcement of the mand requires having the right reinforcer at hand.

The goal of a tact intervention program is to establish a fluent and generative tact repertoire that is functional for the child, and integrated with the other speaker and listener repertoires. A tact program for a child depends on the results of an assessment of her existing tact repertoire, along with the other verbal operants and related skills, as well as possible barriers to tact acquisition (Sundberg, 2014). Once these data are obtained, the tact classification system, the different parts of speech, and ABA teaching procedures can provide a framework and guide for an intervention program. The response form for tacts can be spoken words, signs, written stimuli (including Braille), and, under some circumstances, icon selection. Tacts should also be established across the different sensory receptors affected by nonverbal stimuli (e.g., tactile, gustatory). A number of language curricula are available that contain the core elements of a tact intervention program (e.g., Barbera & Rasmussen 2007; Leaf & McEachin, 1999; Lovaas, 2003; Maurice et al., 1996; Sundberg & Partington, 1998).

Multiple Control

Toddlers begin to emit two-word utterances involving novel combinations of nouns and verbs, and other parts of speech, around 2 years of age (e.g., Brazelton, & Sparrow, 2006). For a child who fails to acquire noun-verb combinations or other types of multiple responses (e.g., adjective-noun), teaching the skill involves arranging multiple forms of nonverbal stimulus control (e.g., a ball rolling down a ramp) while prompting the child to tact the object and its movement (e.g., "ball rolling"), fading prompts, and reinforcing the highest quality responses. These types of multiple tacts, once acquired, should be generalized, extended, and framed with the child's other verbal and nonverbal activities. When a child's verbal repertoire reaches the point where novel multiple antecedent stimuli evoke corresponding multiple responses, a new cusp has been achieved.

True Language

Also around 2 years of age is the developmental period at which "true" language has been said to occur (e.g., Hayes et al., 2001; Hockett, 1960; Horne & Lowe, 1996; Malott, 2003). Interestingly, that is also about the developmental point where C-BiN (Horne & Lowe, 1996), equivalence relations (Sidman, 1994), relational framing (Hayes et al., 2001), and joint control (Lowenkron, 1998) begin to emerge. True language may be the point at which these powerful behavioral cusps begin to demonstrate their generative effects

Tacting Private Events

Teaching a child with language delays to tact stimuli arising within her body can be difficult. A child's collateral behavior is the most common method for identifying the presence of a private event. Facial expressions, pauses, body movement, particular words, and certain phrasing have the power to "give away" a private event such as an eagerness, unwillingness, or apathy. Collateral behavior, such as a child holding her stomach, can be used as a teaching opportunity to establish the tact "sick" because the relevant sources of private control are present, allowing for the use of transfer of stimulus procedures. Public stimuli can also accompany a private event (e.g., blood) and provide the observer with a teaching opportunity.

Duplic and Codic Assessment

Assessing the strength of verbal and nonverbal behavior controlled by verbal S^Ds can be organized by the technical distinctions between the duplic, codic, intraverbal, and listener relations. The duplic repertoires (echoic, motor imitation, copying text) are the simplest of the verbal operants to assess because a finite number of units compose each of these atomic repertoires (Alessi, 1987; Palmer, 2012). For example, measuring the strength of an early echoic repertoire can be accomplished with a one-page form (Esch, 2014). Codic relations (textual and taking dictation) are also easy to assess but involve additional skills such as letter discriminations and phonetics. The intervention section only contains duplic applications. A presentation of the codic applications (reading, writing, and spelling) is beyond the scope of the current chapter.

Duplic Intervention

The ability to duplicate the actions and sounds of others plays a critical role in the acquisition of language, social behavior, and generative learning capabilities (e.g., Baer, Peterson, & Sherman, 1967; Greer & Ross, 2008; Horne & Lowe, 1996; Lovaas, 2003; Lowenkron, 1998; Palmer, 2012; Poulson, 2003). For example, imitating an adult playing peek-a-boo allows for specific back-and-forth interactions between the infant and adult. These interactions can generate direct and automatic reinforcement that further establishes other aspects of an imitative repertoire. A generalized imitative repertoire is an important behavioral cusp because it allows a child to easily learn new behaviors without direct teaching or a reinforcement history (e.g., how to put on socks or connect train tracks). Motor imitation can also have a linguistic function (relating to sign language or gestural communication), and when it does, it can be classified as duplic behavior and serve a verbal function that is parallel to the echoic relation. If duplic behavior fails to develop for an infant-toddler, intervention may be necessary.

Teaching Motor Imitation (Relating to Sign Language)

Many nonvocal children with disabilities have been successful in acquiring sign language as a response form when motor imitation is easier to establish than echoic skills. The first goal of a formal program is to teach the child to copy simple motor behaviors emitted by others. Establishing this repertoire can be accomplished with standard ABA procedures. For children who need further assistance, another method is to place an imitative response within a mand frame, that is, combine the motor S^D with an MO and specific reinforcement for training purposes. For example, if a child demonstrates a strong MO for music, an imitative trial (e.g., tapping an arm as the sign "music"), along with physical prompts, can be conducted while that MO for music is strong, allowing for a momentarily strong form of specific reinforcement to be utilized as a teaching tool (Sundberg & Partington, 1998).

Teaching Echoic Behavior

A child's ability to echo words on command is one of the most important measures of her potential for language acquisition (e.g., Lovaas, 1977). The echoic is valuable because it can be used to establish mand, tact, intraverbal, and listener skills. However, a person's inability to echo words presents a major barrier to future language development, and thus is often a key component of early intervention programs (e.g., Leaf & McEachin, 1999; Lovaas, 2003; Ross & Greer, 2003; Sundberg & Partington, 1998). An echoic repertoire can be established with the use of ABA teaching procedures. However, for those who struggle with learning to echo others, the mand frame procedure described for imitation training can also be used with echoic behavior.

Automatic Reinforcement

Echoic behavior can also be established by pairing a word with existing forms of reinforcement. For example, if just prior to pushing a child on a swing, the word "swing" is emitted by an adult, the word "swing" may eventually become an S^r. Research has shown that this pairing procedure can increase the rate of a child's vocal play and result in the emission of targeted sounds under echoic control (e.g., Esch et al., 2009; Rader et al., 2014; Stock et al., 2008; Sundberg et al., 1996). Children who are having difficulty acquiring an echoic repertoire may benefit from a combination of these procedures, as well as the use of sign language or a selection-based icon system.

Listener Assessment

A listener assessment can be framed by the parts of speech and the five types of verbal discriminations previously described for listeners. In addition, an assessment of self-listener skills is essential (Esch, Esch, McCart, & Petursdottir, 2010; Greer & Ross, 2008). Measures should also be taken to assess a child's ability to serve as an audience, mediate reinforcement, and demonstrate emotional listener responding.

Simple Verbal Discriminations

The assessment of the strength of simple verbal discriminations for a listener involves determining if verbal stimuli reliably evoke a corresponding nonverbal response. For example, when asked to "walk" the child walks, and when asked to "run" the child runs.

Auditory Conditional Discriminations

An auditory conditional discrimination is more complex because it involves multiple stimuli and multiple responses. The stimuli include both a verbal S^D (e.g., "Touch shoe") and a nonverbal S^D (e.g., a picture of a shoe) in an array of other stimuli. The responses include both scanning the array of nonverbal stimuli and emitting a nonverbal section response (e.g., pointing to the shoe). The assessment of this repertoire would involve sampling various combinations of the parts of speech in increasing levels of verbal and nonverbal complexity (e.g., Dunn & Dunn, 2007; Sundberg, 2014).

Compound Verbal Discriminations

The assessment of listener behavior controlled by compound verbal stimuli involves determining if multiple words collectively generate a new S^D in a listener discrimination task. For example, if the target is to teach a child to discriminate between red and green vegetables and red and green fruits presented in an array, first test each element individually (e.g., "Can you find some red things?" "Can you find some fruits?"), then present combinations of the target verbal stimuli (e.g., "Can you find a green vegetable?) to determine if the compounding of the stimuli evokes a more specific selection response.

Verbal Conditional Discriminations

In a VC^D involving listener behavior, two separate conditional discriminations are required. The first occurs in the antecedent event where one verbal stimulus alters the evocative effects of another verbal stimulus in the same antecedent configuration. For example, in a listener discrimination involving the verbal stimulus "Which one can't you wear?" and an array of nonverbal stimuli (e.g., shirt, ball, computer), the reinforcement history related to the word "can't" alters the discriminative function of the word "wear," allowing it to function as a new S^D that participates in a second conditional discrimination involving the comparison array. In that second conditional discrimination, the shirt is established as an S^Δ and the ball and computer are established as S^Ds. Contrasting "can't" with "can" embedded in the phrase will determine if a VC^D is occurring. The assessment of a listener's VC^D repertoire can be quite complex, but would involve sampling various combinations of the parts of speech in increasing levels of verbal and nonverbal complexity (e.g., Devine et al., 2016; Smith et al., 2016; Sundberg & C. A. Sundberg, 2011).

Verbal Function-altering Effect

For a listener, the function-altering future effect involves a change in the evocative effect of a nonverbal stimulus encountered at some point in the future (Schlinger & Blakely, 1994). For example, when a child is given a certain instruction by a parent involving a future nonverbal event, such as "When

the commercial comes on, go brush your teeth," the effects of the commercial as a nonverbal S^D can only be determined later when a commercial comes on. The assessment of this repertoire would involve presenting a variety of these types of verbal instructions and nonverbal contexts.

Listener Intervention

If an infant or toddler fails to demonstrate early listener skills, there are a variety of intervention programs and materials that have produced successful results (e.g., Leaf & McEachin, 1999; Lovaas, 2003; Maurice et al., 1996). Early listener discrimination skills should be established simultaneously with tacting skills (e.g., Petursdottir & Carr, 2011). For example, while teaching a child to tact a shoe, also ask the child to select a shoe from an array. This mixed verbal behavior teaching format (Sundberg & Partington, 1998) consists of interspersing speaker and listener teaching trials and can help establish generative C-BiN repertoires (Greer & Ross, 2008; Horne & Lowe, 1996).

Learning to be a listener involves the capability to discriminate between words (or signs, text, etc.), especially when they appear in novel combinations, in different nonverbal environments, and under different MOs. The five types of verbal discriminations can provide some organization to the process of teaching a child with autism or other intellectual disabilities to comprehend words and sentences as a listener.

Simple Verbal Discriminations

Teaching simple verbal discriminations to a listener involves establishing different nonverbal behaviors under the control of different verbal stimuli. For example, teaching a child to clap when someone says "clap" and to jump when someone says "jump" involves a discrimination between words. Establishing these skills can be accomplished with ABA teaching methodology.

Auditory Conditional Discriminations

A common listener training procedure (also termed receptive labeling) involves asking a child to identify objects, actions, properties, locations, and so forth, by emitting some form of selection behavior (e.g., upon hearing "Where's the nurse?" a child selects a picture of a nurse from a comparison array). A critical component of developing auditory conditional discriminations is the configuration of the array containing the target item or activity. Initially, the array should be clear and simple and contain distinct items (e.g., a banana and a cat). Once a child acquires these beginning discriminations, training should systematically progress to more complex verbal stimuli and arrays. As with the tact, the different parts of speech can serve as a classification and curriculum guide for listener conditional discrimination training, and a number of teaching materials and programs are available for developing this repertoire (e.g., Leaf & McEachin, 1999; Lovaas, 2003).

Compound Verbal Discriminations

Establishing compound verbal discriminations for a listener involves the arrangement of different parts of speech into single phrases and the manipulation of corresponding nonverbal arrays (e.g., Devine et al., 2016; Eikeseth & Smith, 2013). Teaching these types of verbal and nonverbal listener discriminations involves presenting the combination of individual words that are already in a child's listener repertoire. For example, a child may be able to respond as a listener to "wiggle" by wiggling, and "fingers" by showing an adult his fingers, but when asked "Can you wiggle your fingers?" the child cannot respond correctly. To teach this skill, imitative prompts can be used and faded, with the goal of establishing "wiggle fingers" as a new compound verbal S^D that evokes finger wiggling. Eventually new nouns and verbs can be added to the instruction (e.g., wiggle toes, point your finger). Matrix training provides an efficacious and generative format for establishing these types of discriminations.

Verbal Conditional Discriminations

Teaching a child to emit VC^Ds as a listener can be accomplished with a procedure identified as listener responding by function, feature, and class (LRFFC) (DeSouza et al., 2019; Smith et al., 2016; Sundberg, 2014; Sundberg & Partington, 1998). In LRFFC training, a child is asked to identify nonverbal stimuli when given verbal S^Ds based on the things that one can do with an object (e.g., ride it, wear it), its features or attributes (e.g., wheels, sleeves), or its class or category (e.g., vehicles, clothing). Elements of these different verbal stimuli can be combined and contrasted to establish the necessary VC^Ds. For example, given a scene of a farm, an adult asks a child a series of LRFFC questions such as "Which animal eats hay?," "Who grows the hay?," "Which animal lays eggs?," and "Who cooks eggs?" The different parts of speech can be used as a framework for LRFFC training procedures. A major advantage of this type of listener training is that it establishes VC^Ds that can also be used to teach intraverbal behavior.

Verbal Function-altering Effect

Teaching a child to respond to verbal stimuli that change the effects of future nonverbal stimuli can be accomplished by gradually increasing the delay between the initial verbal stimulus and contact with the targeted nonverbal event. For example, if the target relation involves telling a child, "When your father gets here show him your award," training could begin with a brief delay between the instruction and the arrival of the father, and then gradually increase the delay. Once acquired, new instructions along with new nonverbal contexts should be introduced to develop this advanced form of listener behavior.

Intraverbal Assessment

Intraverbal relations are perhaps the most difficult of the elementary verbal operants to assess and quantify. Although early intraverbal behavior emitted by children is usually straightforward, intraverbal relations quickly become more complex as multiple verbal stimuli and multiple responses participate in verbal episodes with listeners. The four types of verbal discriminations, along with the parts of speech, and multiple control involving nonverbal S^Ds and MOs, can provide a framework for intraverbal assessment (Devine et al., 2016; Sundberg, 2014, 2016a; Sundberg & C. A. Sundberg, 2011). In addition, an assessment of I-BiN repertoires can identify the strength and effectiveness of this higher order verbal operant (Greer & Ross, 2008; Miguel, 2016).

Simple Verbal Discriminations

Many examples of simple intraverbal discriminations can be sampled during an assessment (e.g., songs, fill-in-the-blank games, common associations). Performance on these types of verbal tasks can be compared to the performances of typically developing children for a developmental perspective.[5]

Compound Verbal Discriminations

The assessment of a speaker's intraverbal behavior controlled by a compound verbal stimulus is similar to the process for the listener, except without the presence of nonverbal stimuli. For example, when the question "Can you tell me a red fruit?" is presented without any pictures and a child says "apple," and when asked, "Can you tell me a red vegetable?" she says "pepper," this constitutes a measure of intraverbal behavior controlled by a compound verbal S^D.

Verbal Conditional Discriminations

The assessment process for intraverbal behavior controlled by verbal conditional discriminations is similar to the process for the listener, except the strength of responding is tested in the absence of nonverbal stimuli. Assessing this repertoire involves the use of contrasting verbal stimuli inserted into intraverbal autoclitic frames. For example, ask a child questions that require conditional intraverbal responding, such as "What can you wear on your feet" along with "What can you wear on your head," "What grows on your head," "What grows in a garden?" and so on.

Verbal Function-altering Effect

Verbal S^Ds can also transform the functional effects of verbal stimuli encountered in the future. The assessment of a child's ability to demonstrate these effects requires the targeted intraverbal to be observed later in time. For example, give the child a verbal instruction such as "When your uncle asks you about your award, tell him you studied hard for it." Then, when the uncle comes in and asks the targeted question, score the child's response. The assessment of this repertoire would involve presenting a variety of these types of verbal instructions and future measures of their effects.

Intraverbal Intervention

Prior to acquiring an intraverbal repertoire, a toddler has already acquired a sizable collection of mands, tacts, and listener skills. An intraverbal repertoire is typically the last of the initial verbal operants acquired (codic skills are acquired later). Some of the first intraverbal relations are simple phrases, songs, fun fill-ins, and common associations. A child's intraverbal repertoire usually expands rapidly between the ages of 2 and 3. This growth corresponds to increases in the complexity of the tact and listener discrimination repertoires, thus providing a child with content to intraverbally talk about. The four types of verbal discrimination (Figure 18.4) and the parts of speech provide organization for early intraverbal training. In addition, the development of intraverbal behaviors by typically developing children can offer a frame of reference for an intraverbal program (Devine et al., 2016; Kisamore et al., 2016; Sundberg, 2014; Sundberg & C. A. Sundberg, 2011).

Simple Verbal Discriminations

Teaching beginning intraverbal behavior to a child involves establishing verbal behavior under the control of simple verbal stimuli. For example, an adult who is playing peek-a-boo may say "peek a . . . " and then implement a brief delay before saying "boo" to evoke the intraverbal response "boo" from the child. Fill-in-the-blank phrases ("You ride a . . . ") are usually easier than WH questions ("What can you ride?") because fill-ins often just involve a simple verbal discrimination, whereas WH questions ultimately involve verbal conditional discriminations. The primary goal at this early intraverbal level is to break the targeted verbal responses free from echoic control, while teaching the child to verbally discriminate among the many verbal stimuli she may encounter. In addition, the fill-in to WH question procedure helps to establish early I-BiN repertoires (Greer & Ross, 2008; Horne & Lowe, 1996).

Compound Verbal Discriminations

Teaching intraverbal behavior controlled by compound verbal stimuli is similar to the procedure previously described for a listener discrimination, except a verbal response is required. For example, the words "brother" and "name" may independently evoke certain behaviors, but when combined to "brother's name" a new stimulus occurs and by prompting, fading, and reinforcing a specific response under these circumstances (e.g., "Cory") compound verbal stimulus control can be established.

Verbal Conditional Discriminations

When considering an intervention program targeting the establishment of VC^Ds, it is important to ensure that the child demonstrates the prerequisite speaker and listener skills (DeSouza et al., 2019; Sundberg & C. A. Sundberg, 2011). For example, prior to being able to intraverbally talk about hot and cold, it is essential that the tact and listener repertoires (e.g., LRFFC) regarding hot and cold are established. If a child can demonstrate these prerequisite skills, and this repertoire has been targeted as a priority, one can turn to a sizable body of research and clinical applications on intraverbal responses controlled by multiple verbal stimuli (for reviews see Aguirre, Valentino, & LeBlanc, 2016; Petursdottir, 2018).

For example, Kisamore et al. (2016) demonstrated the effects of three different procedures for establishing multiply controlled intraverbals with seven participants with autism. The first procedure involved a standard prompt delay and error correction, but this procedure was effective for only three of the participants. Two of the participants benefited from the requirement to echo the question first, termed differential observing response (DOR), and two participants required other modifications of the procedures. Figure 18.7 shows the results for the two participants who failed with prompt delay and error correction, but benefited from the echoic DOR procedure. The data for both participants showed that intraverbal responding during the prompt delay and error correction was variable and did not meet criterion.

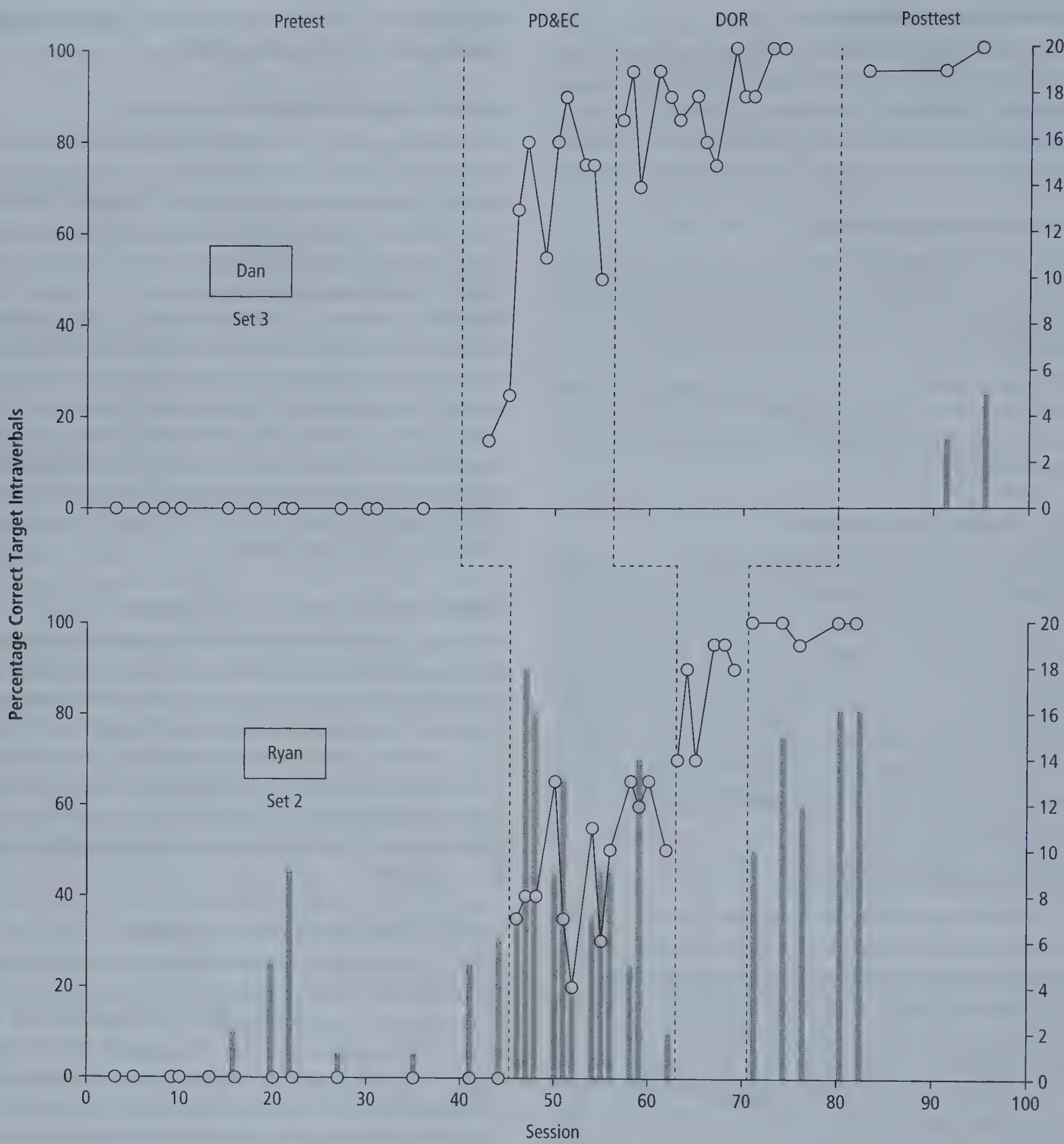


Figure 18.7 Results for the two participants who failed with prompt delay and error correction (PD&EC), but benefited from the differential observing response (DOR) procedure. The open circles show acquired intraverbal responses across conditions, and the bars in the right panel show the number of errors across conditions.

From Teaching Multiply Controlled Intraverbals to Children And Adolescents with Autism Spectrum Disorders" by A. N. Kisamore, A. M. Karsten, & C. C. Mann, © 2016. Reproduce with permission of John Wiley & Sons Inc.,

However, with introduction of the DOR procedure, responding increased to mastery performance.

Verbal Function-altering Effect

The goal in establishing delayed intraverbal behavior is to teach a child how to respond to specific verbal S^Ds that might occur later in time. An instructional program would involve known intraverbal relations and altering the temporal distance between the first and second presentation of the target verbal S^D. For example, if the goal is a future intraverbal interaction in a restaurant, initial conditioning might involve telling a child, "When the waiter asks you, 'What do you want to order?' you tell him chicken nuggets", then engage in role playing involving a second adult and gradually increase the time between each adult's verbal S^D presentation.

Autoclitic Assessment

The following prerequisite skills are necessary to benefit from autoclitic instruction: (1) A speaker's existing verbal skills are sufficient to support secondary verbal behavior,

(2) a speaker demonstrates self-listening skills, and (3) a speaker can tact aspects of her own verbal and nonverbal behavior. Developmental milestones can be helpful for identifying when these types of repertoires emerge (generally around 24 to 36 months of age).

Autoclitic Mands

Secondary autoclitic mands begin to emerge in a speaker's mand repertoire after primary mands have been acquired. For example, does a child put extra stress on certain words, add tags of immediacy, or pause before saying something? If these secondary verbal behaviors are controlled by some MO regarding how a listener should respond to a primary response, along with the relevant history of specific reinforcement, they can be classified as autoclitic mands.

Autoclitic Tacts

The assessment of autoclitic tacts can reveal if a speaker's behavior is controlled by her own verbal behavior and is modified due to its special effects on listeners (e.g., emitting an autoclitic of strength, such as "I KNOW it is mine," or an autoclitic of weakness, such as "I THINK it is mine"). These types of autoclitic tacts begin to demonstrate capabilities with verbal complexities such as self-editing, covert verbal behavior, and verbal self-awareness (Greer & Ross, 2008; Skinner, 1957, 1974). A child's ability to emit tacts controlled by these various types of nonverbal S^Ds can be compared to that of typically developing children for a developmental perspective.

Autoclitic Intervention

Establishing an autoclitic repertoire can be difficult due to the private nature of the relevant controlling variables. In addition, all autoclitic relations involve multiple controlling variables, multiple responses, and a speaker serving as her own listener. Teaching autoclitic mands involves capturing or creating the relevant MOs and using ABA teaching procedures to establish a secondary mand relation. For example, in social interaction with peers a louder voice or a certain phrase may function as an autoclitic mand to prompt a listener to attend better to what the speaker says. Establishing these types of autoclitics can be accomplished with basic ABA procedures, including modeling and role playing.

Teaching an autoclitic tact is complicated by the fact that the target nonverbal S^D must originate from some aspect of the child's own words, and it is that S^D that should evoke the child's autoclitic response. The task for the adult is to bring the child's verbal behavior under the control of some specific aspect of the child's own ongoing verbal behavior. This is not an easy task. For example, a child must first be emitting primary responses in order to generate stimuli to secondarily tact, but a child's own words may be transient, private, multiple, or completely missed by the child. They also may have to compete with other words, ongoing listener behavior, current MOs, and any number of variables that may emerge in verbal exchanges. Two procedures for establishing autoclitic tacts will be described. As with the autoclitic mand, capturing or creating the relevant controlling variables can provide the necessary elements to establish autoclitic tacts.

Capturing Autoclitic Tact Teaching Opportunities in the Natural Environment

Once a repertoire of primary verbal operants is established, autoclitic tacts can emerge in typically developing children without formal instruction. This is because listeners, often unintentionally, provide differential reinforcement for autoclitic behavior. Listeners also prompt speakers for autoclitic expansions of their primary response (e.g., "Where did you hear that?"). Certain settings and activities may be more conducive to evoking primary verbal behavior, such as play activities, arts and crafts, or outdoor events and games. Often, autoclitic behavior is already occurring in the context of ongoing verbal behavior, and if an adult can identify its presence, the adult can provide reinforcement for that behavior. For example, upon observing a wave in a water pool, a child emits the response, "That was SO big"—a response that appears to contain a descriptive autoclitic. An adult can reinforce the additional autoclitic behavior by drawing attention to the "SO" part (e.g., "Wow that was the biggest one!").

Formal Arrangement of Autoclitic Tact Contingencies

If a child demonstrates the prerequisite skills identified above, she may be prepared to learn to emit autoclitic behavior through direct teaching procedures (e.g., multiple exemplar instruction). For example, if the targeted autoclitic relation involves a child informing the listener of the sense mode that is affecting her primary tact, she must first learn to discriminate between seeing things, hearing things, smelling things, and so on. A formal teaching procedure could begin with a primary response evoked with a standard tact trial to set up the necessary conditions for tacting the nonverbal source of control for that response (e.g., visual or auditory). When the stimulus is visual, the echoic prompt "I see the ball" can be used to prompt a response (at this point "I see" is not autoclitic; it is a primary echoic operant). When the stimulus is auditory (e.g., the sound of a ball bouncing), the echoic prompt "I hear the ball" can be provided. Both types of echoic prompts need to be faded, trials presented randomly, and stimulus control transferred to the type of sensory antecedent that controls the primary tact "ball." It is important to transfer what is acquired through direct instruction to natural environment contingencies where sensory information is valuable to the listener (e.g., hide-and-seek games, emergency situations). Similar procedures could be conducted for the many other types of descriptive autoclitics, as well as for quantifying and relational autoclitics.

Autoclitics are an important aspect of a verbal repertoire, but the private nature of the controlling contingencies make them difficult to study. Fortunately, a growing body of autoclitic research is available (e.g., Hübner, Austin, & Miguel, 2008; Lowenkron, 1998, 2006; Luke et al., 2011; Palmer, 2011; Sidener & Michael, 2006; Speckman et al., 2012; C. T. Sundberg et al., 2018).

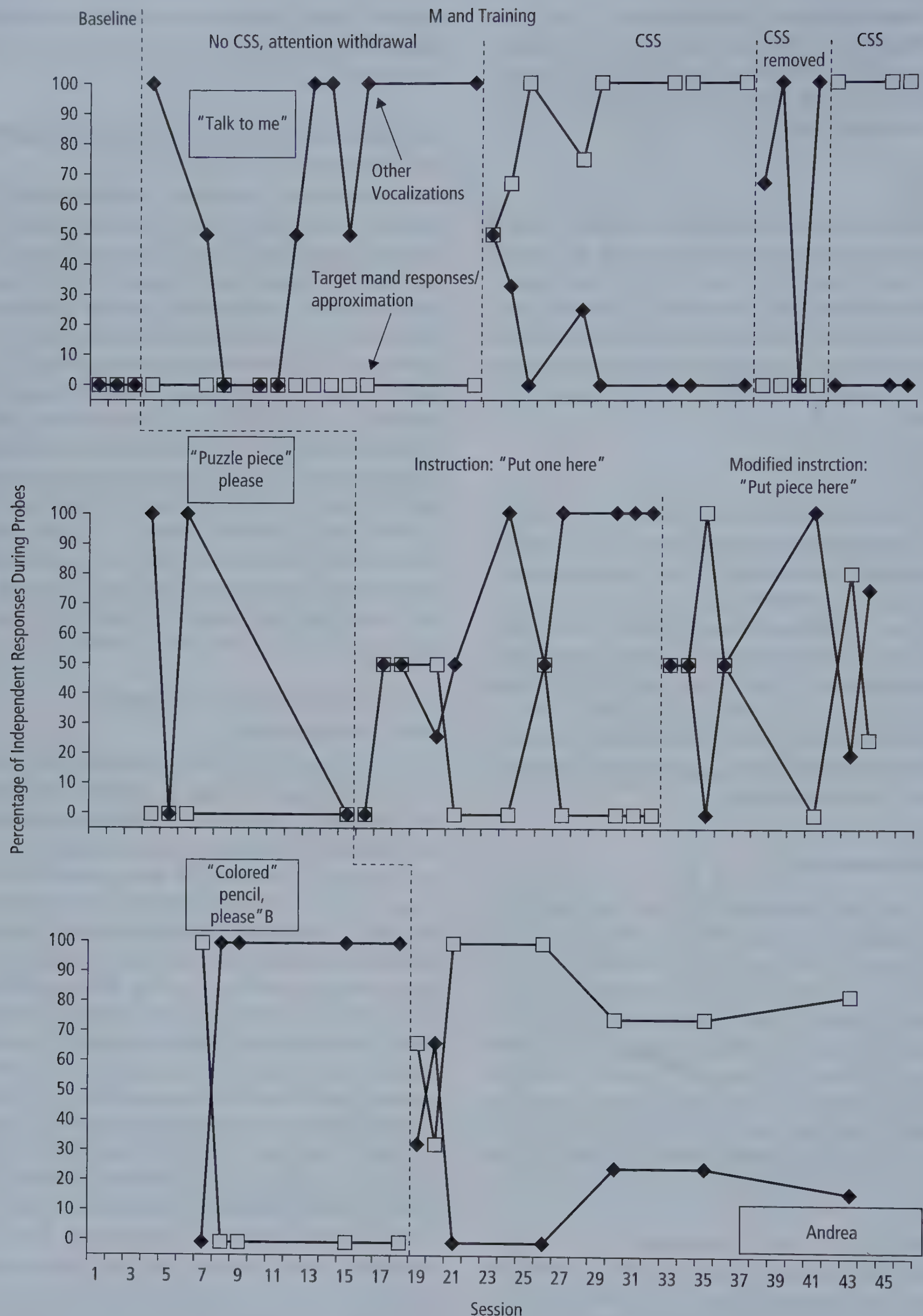


Figure 18.8 The results for Andrea show her absent or weak mands during baseline; the open squares represent the specific target words (e.g., "colored pencil"), and the closed diamonds represent indirect vocal words as mands (e.g., "I don't know what to do," "OK"). The first condition was ineffective in evoking the target mand, but the second condition involving various types of verbal prompts such as a contingency specifying stimulus (CSS), produced successful manding and reduced inappropriate mands.

THE LOSS OF VERBAL BEHAVIOR

Skinner's analysis of VB can serve as a guide for language assessment and intervention for an individual experiencing a verbal loss, much in the same way it does for initial language acquisition. Verbal repertoires can weaken significantly, as observed in cases of dementia, Alzheimer's disease, aphasia, and traumatic brain injury (Baker, LeBlanc, & Raetz, 2008; Gross, Fuqua, & Merritt, 2013; Palmer, 1991; Sundberg, 2016b; Sundberg, San Juan, Dawdy, & Arguelles, 1990; Trahan, Donaldson, McNabney, & Kahng, 2014). An adult with dementia may, for example, experience memory loss, engage in off-topic verbalizations, or neglect important aspects of daily living (e.g., Bourgeois, 1993; LeBlanc, Raetz, & Feliciano, 2011).

In his discussion of aphasia and the loss of verbal behavior, Skinner (1957) points out the following: "The order of damage seems to follow the order of 'difficulty' deducible from the availability of a minimal repertoire. Textual and echoic behavior often survive (unless relevant sensory defects are involved) while intraverbals and tacts appear to be most vulnerable" (p. 219). Identifying the strength of each verbal operant and the functional sources of control for an adult's impaired verbal behavior can provide direction for an intervention program. For example, data regarding the nature of a person's loss can be obtained by assessing intraverbal repertoires regarding recent and past events, mands for information, tacting complex stimuli, and the occurrences of odd or irrelevant VB (e.g., Gross et al., 2013; Sundberg, 2014).

The body of research on applying Skinner's analysis of verbal behavior to problems related to the loss of language is expanding (e.g., Dixon, Baker, & Sadowski, 2011; LeBlanc et al., 2011; Oleson & Baker, 2014). For example, Oleson & Baker (2014) examined manding problems experienced by older adults with dementia. These authors suggested that "[t]raining mands could not only lead to an improved quality of life by providing individuals with dementia the ability to access desired items and activities . . . but may also reduce aggression and other challenging behavior" (p. 115). In Oleson and Baker's (2014) study, two participants with dementia experienced a mand training condition consisting of contriving an MO (e.g., providing paper and no colored pencils), verbal prompts, fading, specific reinforcement, and error correction. The results revealed that the procedure was effective with one participant, but results were inconsistent with the second participant. Figure 18.8 shows the results for the successful participant, Andrea, who acquired two mands and partially acquired a third. In the mand baseline condition, Andrea emitted indirect vocal words as mands (e.g., "I don't know what to do," "OK"), as indicated by the closed triangles, but rarely emitted the specific target words (e.g., "colored pencil"), as indicated by the open squares. The first condition consisted of contriving the EO alone, which was ineffective in evoking the target mand. However, the second condition involving the contrived MO plus an indirect supplemental verbal prompt resulted in successful manding and reduced inappropriate mands. These data are similar to the results reported for Jason in the opening vignette, where Jason's inappropriate mand was replaced by an acceptable response form through a basic mand training procedure. Many other teaching procedures have proven effective for children with autism or other intellectual disabilities that can be modified and used for individuals experiencing a loss of verbal behavior.

SUMMARY

Skinner's (1957) Analysis of Verbal Behavior

1. Skinner argues that language constitutes learned speaker and listener behavior that is acquired, maintained, and extended by environmental contingencies of reinforcement.
2. A given utterance is comprised of formal properties (the words spoken) and functional properties (why those words were spoken).
3. The formal properties of language involve the topography and structure of verbal responses usually classified in terms of the parts of speech, clauses, and sentences.
4. The functional properties of language involve the causes of a response often referred to as semantics, or the meaning of words.
5. From a behavioral perspective, the meaning of a word can be found in the immediate and historical antecedents and consequence that evoke that word.
6. By treating language as learned behavior emitted by speakers and listeners, the focus of the analysis and its applications is on the contingencies responsible for that behavior.
7. Skinner (1957) defined verbal behavior "as behavior reinforced through the mediation of other persons" (p. 2), but those other persons "must be responding in ways which have been conditioned *precisely in order to reinforce the behavior of the speaker*" (p. 225).
8. Language involves an interaction between a speaker and a listener, including when a speaker is her own listener. Skinner terms the basic unit of analysis for this interaction a *verbal episode*.
9. In a verbal episode, a speaker emits any type of verbal behavior (e.g., mand, tact, intraverbal), in any form (e.g., speech, sign language, eye contact).
10. In a verbal episode, a listener (1) serves as an audience for a speaker, (2) provides reinforcement for a speaker, and (3) responds in specific ways to the speaker's behavior.

11. The roles of speaker and listener alternate in verbal episodes, and usually involve covert speaker and listener behavior as well.

12. Verbal behavior is not restricted to spoken words. Any form of operant behavior can acquire a verbal function (e.g., sign language, fingerspelling, icon selection, Braille, facial expressions, intonation).

13. In topography-based verbal behavior, each verbal relation involves a distinguishable response form given some specific controlling variable.

14. In selection-based verbal behavior, the response form is the same for each verbal relation (e.g., pointing, selecting); what is conveyed to the listener is the stimulus indicated.

15. A behavioral cusp is a form of generative learning in which the acquisition of one skill or specific set of skills enables the acquisition of other skills without direct teaching or a history of reinforcement. A cusp can allow access to new MOs, stimuli, reinforcers, and punishers, which in turn generate new behavior.

16. The basic components of a speaker's verbal repertoire are called the elementary verbal operants (e.g., mand, tact, intraverbal).

17. The mand is a verbal operant in which the form of the response is under the functional control of motivating operations (MOs) and a history of specific reinforcement.

18. The tact is a verbal operant in which the form of the response is under the functional control of a nonverbal S^D and a history of generalized S^r.

19. There are three types of dupic relations: echoic, motor imitation, and copying text, and two types of codic relations: textual and taking dictation.

20. The echoic (duplic) is a verbal operant in which the form of the response is under the functional control of a verbal S^D that has formal similarity with a verbal response and a history of generalized S^r.

21. Formal similarity occurs when the controlling antecedent stimulus and the response product (a) share the same sense mode (i.e., both stimulus and response are visual, auditory, or tactile) and (b) physically resemble each other.

22. The intraverbal is a verbal operant in which the form of the response is under the functional control of a verbal S^D that does *not* have point-to-point correspondence to the verbal response, but does have a history of generalized S^r.

23. Point-to-point correspondence between the stimulus and the response or response product occurs when the beginning, middle, and end of a verbal stimulus matches the beginning, middle, and end of a verbal response.

24. The textual relation (codic) is a verbal operant in which the form of the response is under the functional control of a written verbal S^D that does not have formal similarity, but does have point-to-point correspondence and a history of generalized S^r.

25. Taking dictation (codic) is a verbal operant in which the form of the response is under the functional control of a verbal S^D that controls a written, typed, Braille punched, or Fingerspelled response. Like the textual relation, there is no formal similarity, but there is point-to-point correspondence and a history of generalized S^r.

26. Listener discriminations (commonly referred to as receptive language) can be defined as a verbal S^D that evokes a corresponding nonverbal response (or class of responses) due to a history of generalized S^r.

27. The generalized ability to match nonverbal sample stimuli to nonverbal comparison stimuli can facilitate the establishment of equivalent and nonequivalent relations.

28. Classifying verbal relations allows us to more precisely identify, organize, and account for the specific variables that might be responsible for a given utterance.

The Verbal Operants and Listener Behavior in more Detail

29. Mands can be classified by identifying the different types of MOs that control manding behavior (e.g., UMOs, CMOs), and by the different ways mands affect the behavior of the listener and produce specific reinforcement.

30. The defining consequences for the mand constitute what Skinner identifies as specific reinforcement. Specific reinforcement involves reinforcement that is established by its relation to a specific MO (e.g., water deprivation establishes water as specific reinforcement).

31. Several types of MOs are related to aversive stimulation (e.g., too hot, loud noises, demands), and their removal functions as specific negative reinforcement.

32. When sufficient specific positive or negative reinforcement has been provided, it abolishes the MO (e.g., satiation).

33. Any instance of human behavior is usually a function of multiple variables operating simultaneously. There are two types of multiple control: convergent multiple control and divergent multiple control.

34. In convergent multiple control, a single response is controlled by more than one antecedent variable.

35. In divergent multiple control, a single antecedent variable controls more than one response.

36. It is common to have multiple MOs and S^Ds operating simultaneously within a single verbal episode.

37. The involvement of multiple antecedent variables can enhance or weaken an evocative effect.

38. A private event is only felt by one person. However, private events can have immediate behavioral effects (e.g., MOs, covert verbal S^Ds), and should be taken into account for a complete interpretation of behavior.

39. Public accompaniment occurs when a publicly observable stimulus reliably occurs with a private stimulus.

40. Collateral responses are publicly observable behaviors that reliably occur with a private stimulus.

41. Once a child has acquired basic echoic, mand, tact, and listener repertoires, the combination of these cusps can produce emergent mands and generative learning in a number of different ways (e.g., tact to mand transfer, incidental mand acquisition).

42. Classifying tacts can be accomplished by identifying the different types of nonverbal S^Ds that control responses, and by the different ways these verbal responses affect the behavior of the listener.

43. Nonverbal S^Ds commonly participate in verbal episodes with other types of antecedent variables (e.g., CMO-Ts, verbal S^Ds).

44. Once a tact relation has been established, the tact response can extend to novel stimulus conditions through the process of stimulus generalization.

45. In generic tact extension, the novel stimulus shares all of the relevant or defining features associated with the original stimulus.

46. In metaphorical tact extension, the novel stimulus shares some, but not all, of the relevant features of the original stimulus.

47. In metonymical tact extension, the novel stimulus shares none of the relevant features of the original stimulus configuration, but some irrelevant but related feature has acquired stimulus control.

48. In solecistic tact extension, a stimulus property has no clear connection to the original stimulus, but the novel stimulus has acquired control through some type of reinforcement history.

49. Listeners provide five different types of generalized S^r for tacting: (1) educational reinforcement, (2) escape from or avoidance of aversive stimulation, (3) extending the listener's contact with the environment, (4) automatic reinforcement, and (5) a history of contiguous or correlated usage.

50. Skinner (1957) used the term "automatic" to identify circumstances in which behavior is evoked, shaped, maintained, strengthened, or weakened by environmental variables occurring without direct manipulation by other people.

51. Automatic reinforcement can occur when the reinforcement is provided by the response product emanating from the behavior.

52. Automatic reinforcement can occur when the reinforcement is provided by the behavior's reinforcing effects emanating from the physical environment.

53. Generative verbal learning is a behavioral effect where previously acquired speaker and listener skills enable or accelerate the acquisition of other speaker and listener skills, without dependence on direct teaching or a history of reinforcement.

54. Once a child has acquired basic echoic, tact, listener, and matching-to-sample repertoires, the combination of these cusps can produce emergent tacts in a number of different ways (e.g., stimulus equivalence, recombinative generalization, common bidirectional naming, joint control, and relational framing).

55. Stimulus equivalence occurs when, for example, the direct teaching of an AB relation between a word (stimulus A) and an object (stimulus B), and the direct teaching of a BC relation between the same object (stimulus B) and a different stimulus (stimulus C), can generate several untrained speaker and listener relations (BA/CB symmetry and AC/CA transitivity). When emergence occurs with both symmetry and transitivity, equivalence has occurred (Chapter 19).

56. Recombinative generalization is a type of multiple control involving learning to respond to novel combinations of antecedent variables as both a speaker and a listener.

57. Bidirectional naming (BiN) is a higher order verbal cusp consisting of the fusing of the speaker and listener repertoires in bidirectional relations. In common BiN, a new word acquired as listener can generate a tact relation without further training, and a new word acquired as a tact can generate a listener relation without further training.

58. Joint control involves the combination of verbal skills (e.g., echoic and tact) and occurs when two separately established antecedents that evoke the same response topography arise simultaneously and produce an emergent S^D that evokes behavior.

59. Relational framing involves an extension of Sidman's equivalence theory to nonequivalent stimulus relations such as comparison, opposition, and difference (Chapter 20).

60. Verbal S^Ds have the same causal status as nonverbal S^Ds in that both types of S^Ds acquire discriminative control over behavior through a history of differential reinforcement.

61. Six types of elementary verbal operants are controlled by verbal S^Ds: echoic, motor imitation (relating to sign language), copying text, textual, taking dictation, and intraverbal. All of these relations are duplic or codic, except for the intraverbal.

62. Intraverbal relations can be classified by identifying the type of verbal stimulus control that evokes the intraverbal response.

63. The intraverbal response can be controlled by four different types of verbal stimulus control: (1) simple verbal discriminations, (2) compound verbal discriminations, (3) verbal conditional discriminations, and (4) verbal functional-altering discriminations.

64. A simple verbal discrimination involves a single-component word or phrase that evokes an intraverbal response.

65. A compound verbal discrimination involves two or more verbal S^Ds (convergent multiple control) that each

independently evoke behavior, but when they both occur in the same antecedent configuration, a different S^D is generated, and a more specific intraverbal response is evoked.

66. A verbal conditional discrimination is a type of convergent multiple control involving a verbal stimulus that alters the evocative effects of another verbal stimulus in the same antecedent configuration, and evokes a corresponding intraverbal response.

67. The verbal function-altering effect occurs when verbal stimuli condition the behavior of a listener by changing the functions of other stimuli and MOs that the listener encounters immediately or later in time (e.g., instructions).

68. Generalized S^r provided by listeners (e.g., praise, approval) establishes verbal stimulus control over verbal behavior, in the same way it does for nonverbal stimulus control.

69. Research has demonstrated that intraverbal relations can emerge from previously acquired tact, intraverbal, and listener repertoires, and even from direct observation. Intraverbal behavior can also emerge through the process of bidirectional naming.

70. Intraverbal bidirectional naming (I-BiN) involves a bidirectional relation that is established between at least two verbal stimuli (e.g., "cat" & "animal"). I-BiN occurs when AB training on a unidirectional intraverbal relation automatically produces a reverse BA intraverbal relation, and vice versa.

Listener Behavior

71. A listener is an active participant in verbal exchanges and serves many roles in this capacity, such as acting as an audience, providing reinforcement for speakers, and taking action upon a speaker's words.

72. The same four types of verbal discriminations described for intraverbal relations can be used as a framework to identify various types of nonverbal listener discriminations. In addition, there is a fifth type of verbal discrimination that is specific to listener behavior commonly termed an auditory conditional discrimination.

73. In an auditory conditional discrimination, a verbal stimulus alters the evocative effect of a nonverbal stimulus due to a history of reinforcement and evokes a corresponding nonverbal response.

74. A listener often engages in overt or covert verbal behavior in the process of listening to a speaker; when this occurs the listener is now a speaker, and his behavior should be classified as such.

75. A speaker can also be her own listener. We can, and often do, talk to ourselves.

Autoclitic Verbal Behavior

76. A speaker often adds autoclitic responses to her on-going primary verbal behavior because autoclitic responses provide a listener with more information about a speaker's primary response, and a listener reinforces a speaker for this extra information.

77. A primary verbal response constitutes any of the verbal operants (e.g., mand, tact, intraverbal).

78. A secondary autoclitic response is controlled by some aspect of the primary verbal operant or by its controlling variables. There are three types of secondary autoclitic behavior: autoclitic mand, autoclitic tact, and autoclitic intraverbal frames.

79. The autoclitic mand enjoins a listener to take some specific action regarding the primary response, and the autoclitic behavior is reinforced by that action.

80. The autoclitic tact is controlled by some nonverbal feature of the primary response or its controlling variables, and the autoclitic response informs the listener of that feature. This behavior is reinforced by listeners who provide generalized S^r.

81. Intraverbal autoclitic frames provide order, agreement, grouping, and composition of larger units of verbal behavior. An autoclitic frame facilitates the connections between words in a given utterance and is reinforced by listeners who provide generalized S^r.

Applications of Skinner's (1957) Analysis of Verbal Behavior

82. Behavior analysis provides three major contributions to language assessment and intervention programs for individuals with special needs: (1) applied behavior analysis technology, (2) a behavioral analysis of human development, and (3) a behavioral analysis of language.

83. Applied behavior analysis can provide professionals and parents with the necessary intervention procedures and scientific methods to teach language skills.

84. The development of language and social skills by typically developing children can provide a valuable guide for a language intervention curriculum.

85. Skinner's analysis of verbal behavior can serve as a framework for specific language assessment and intervention programs.

86. A language assessment for a person with language delays has three primary goals: (1) to identify the nature of the delay or problem, (2) to compare the person's performance with expected norms, and (3) to provide guidance for an appropriate intervention program if necessary.

87. Language assessment for a person with language delays involves measuring the strength of the elementary verbal operants, listener skills, self-listener skills, multiply controlled responding, speaker-listener interactions.

88. Assessing a child's mand repertoire involves identifying the degree to which the various MOs control verbal responses.

89. There are two general methods are used for identifying and controlling the value change of MOs for mand training: (1) capturing the MO value change as it occurs naturally and (2) creating the MO by generating some change that alters the value of a consequence.

90. Assessing a tact repertoire involves determining the degree to which nonverbal stimuli control verbal behavior.

91. The goal of a tact intervention program is to establish a fluent and generative tact repertoire that is functional for the child, and integrated with the other speaker and listener repertoires.

92. A listener assessment can be framed by the parts of speech and the five types of verbal discriminations previously described. In addition, an assessment of a child's self-listener skills is important.

93. Early listener discrimination skills should be established simultaneously with early tacting skills.

94. Learning to be a listener involves the capability to discriminate between words (or signs, text, etc.), especially when they appear in novel combinations, in different nonverbal environments, and under different MOs.

95. The four types of verbal discriminations, along with the parts of speech, and multiple control involving nonverbal S^Ds and MOs, can provide a framework for intraverbal assessment and intervention.

96. The development of intraverbal behaviors by typically developing children can offer a frame of reference for an intraverbal intervention program.

97. There are several prerequisite skills are necessary to benefit from a formal autoclitic intervention program: (1) A speaker's existing verbal skills are sufficient to support secondary verbal behavior, (2) a speaker demonstrates self-listening skills, and (3) a speaker can tact aspects of her own verbal and nonverbal behavior.

98. Generative verbal learning is a major goal for a language intervention program. Procedures from stimulus equivalence, relational framing, bidirectional naming, and joint control can all help to facilitate the emergence of generative repertoires for a person with language delays.

The Loss of Verbal Behavior

99. Verbal repertoires can also deteriorate, as observed in cases of dementia, Alzheimer's disease, aphasia, and traumatic brain injury. Skinner's analysis of VB can serve as a guide for language assessment and intervention for these individuals much in the same way it does for those learning to communicate for the first time.

KEY TERMS

audience
autoclitic
automatic punishment
automatic reinforcement
convergent multiple control
copying a text
divergent multiple control
echoic
formal similarity
generic (tact) extension
impure tact
intraverbal
listener
mand
metaphorical (tact) extension
metonymical (tact) extension
multiple control
point-to-point correspondence
private events
solistic (tact) extension
speaker
tact
textual
taking dictation
verbal behavior
verbal operant

MULTIPLE-CHOICE QUESTIONS

1. The formal properties of language involve the:
   a. Cause of the verbal response
   b. Topography of the verbal response
   c. Unseen forces of verbal language
   d. Language properties are not formal
   Hint: (See "Form and Function of Language")

2. The functional properties of language involve:
   a. Causes of the verbal response
   b. Topography of the verbal response
   c. Language properties do not have functional components.
   d. Unobservable psychic antecedents.
   Hint: (See "Form and Function of Language")

3. A mand is a type of verbal operant in which a speaker:
   a. Differentially responds to the verbal behavior of others
   b. Asks for (or states, demands, implies, etc.) what he or she needs or wants
   c. Repeats the verbal behavior of another speaker
   d. Names things and actions that the speaker has direct contact with through any of the sense modes

   Hint: (See "The Elementary Verbal Operants and Listener Discriminations")

4. An intraverbal response is a type of verbal operant in which a speaker:
   a. Differentially responds to the verbal behavior of others
   b. Asks for (or states, demands, implies, etc.) what he needs or wants
   c. Repeats the verbal behavior of another speaker
   d. Names things and actions that the speaker has direct contact with through any of the sense modes

   Hint: (See "The Elementary Verbal Operants and Listener Discriminations")

5. A tact is a type of verbal operant in which a speaker:
   a. Differentially responds to the verbal behavior of others
   b. Asks for (or states, demands, implies, etc.) what he needs or wants
   c. Repeats the verbal behavior of another speaker
   d. Names things and actions that the speaker has direct contact with through any of the sense modes

   Hint: (See "The Elementary Verbal Operants and Listener Discriminations")

6. A textual operant has
   a. Point-to-point correspondence but no formal similarity
   b. Point-to-point correspondence and formal similarity
   c. No point-to-point correspondence and no formal similarity
   d. None of these

   Hint: (See "The Elementary Verbal Operants and Listener Discriminations")

7. Taking dictation consists of:
   a. A discriminative stimulus evoked by a verbal response that does not have point-to-point correspondence.
   b. Reading without any implications so that the reader understands what is being read
   c. Writing and spelling words that are spoken
   d. None of these

   Hint: (See "The Elementary Verbal Operants and Listener Discriminations")

ESSAY-TYPE QUESTIONS

1. Explain some of the factors in the slow appreciation of Verbal Behavior.

   Hint: (See "Form and Function of Language")

2. Describe the term mand. Be sure to include in your answer the definition of the term and its importance for the early development of language.

   Hint: (See "The Elementary Verbal Operants and Listener Discriminations")

3. Describe the term tact.

   Hint: (See "The Elementary Verbal Operants and Listener Discriminations")

4. Compare and contrast convergent multiple control and divergent multiple control.

   Hint: (See "Multiple Control")

NOTES

1. For ease of reading throughout the chapter, the male pronoun will be used for the listener and the female pronoun will be used for the speaker.
2. However, when using the term *verbal behavior* to refer to Skinner's analysis of language, the "verbal" part involves the behavior of the speaker, but where is the behavior of the listener in that term? Skinner did not neglect the listener in the analysis, but one must add the phrase "and the behavior of the listener" when describing Skinner's analysis of verbal behavior. An alternative is to use the phrasing "speakers and listeners" when describing the complete verbal field.
3. For more detail, see Carbone (2013); Endicott and Higbee (2005); Hall and Sundberg (1987); Landa, Hansen, and Shillingsburg (2017); Lechago, Carr, Grow, Love, and Almason (2010); Oleson and Baker (2014); and Sundberg, Loeb, Hale, and Eigenheer (2002).
4. For reviews and research topics, see Aguirre, Valentino, and LeBlanc (2016); Dymond, O'Hora, Whelan, and O'Donovan (2006); Gamba, Goyos, and Petursdottir (2015); Petursdottir (2018); Petursdottir and Devine (2017); Sautter and LeBlanc (2006); Sundberg (2013).
5. See Sundberg and C. A. Sundberg (2011) for a developmentally sequenced 80-item intraverbal subtest for the *Verbal Behavior Milestones Assessment and Placement Program (VB-MAPP)* (Sundberg, 2014).

PART 8

Developing New Behavior

The five chapters in Part 8 describe methods for developing new behavioral repertoires. Chapter 19, by Carol Pilgrim, outlines the research foundations and core concepts for equivalence-based instruction (EBI) and discusses several key outcomes of EBI, namely, class formation, delayed emergence, class expansion and class merger, transfer of function, and contextual control. The chapter ends with a treatment of how to apply EBI in instructional and generalized settings, as well as applications viewed through the perspectives of Naming Theory and Relational Frame Theory.

Chapter 20, by Tom Critchfield and Ruth Anne Rehfeldt, describes how learning can be engineered with nonequivalent relations. After defining nonequivalence classes and why they matter, these authors present the basic vocabulary for explaining nonequivalence relations, describe types and examples of nonequivalence relations, review key theoretical concepts, address the role of nonequivalence relations in higher-order psychological constructs, and explain how nonequivalence relations provide a basis for enhancing well-being in ways that are quite different from typical approaches in applied behavior analysis.

Chapter 21 details how new behaviors are acquired through modeling, imitation, and observational learning. After defining and providing examples of these terms, and outlining a training protocol for learners who do not display imitative behavior, the chapter describes how practitioners can use modeling as a behavior change strategy and teach children to be skilled observational learners.

Chapter 22 describes how to shape new behavior by reinforcing successive approximations to a terminal behavior. The chapter also includes procedures for increasing the efficiency of shaping and guidelines for using shaping in applied settings.

Chapter 23 explains how discrete responses can be linked to create behavior chains of more complex behaviors. Procedures for task analysis—breaking down complex tasks into sequenced steps for instruction—are described. The chapter addresses the varied methods of chaining and factors affecting the performance of behavior chains.

CHAPTER 19

Equivalence-based Instruction

Carol Pilgrim

LEARNING OBJECTIVES

- Understand the research foundations and core concepts of equivalence-based instructions.
- Learn the designing of equivalence-based instructions procedures.
- Discuss the application and generality of equivalence-based instructions with varied participant populations.
- Discuss the applications stemming from alternative theoretical approaches to relational responding.

It would be difficult to overestimate the role played by symbols in our everyday lives. Consider my commute to work this morning: As I left the house, I discovered a small vase on my vanity, filled by my better half with fresh azaleas from our garden. With considerable goodwill toward the world, then, I pushed the button marked "Start" on my car's dash, moved the shifter to "R," and punched "5" on the console for our local NPR station. An angry debate was in progress on the meaning of "building the wall" for U.S. Democrats and Republicans, for American immigrants (past, present, and future), and, indeed, for observers from around the world. With a quick shift to my shuffled music play list, John Lennon's "Give Peace a Chance" created a powerful contrast. Shortly thereafter, a swing by the drive-through of my local bank branch allowed me to deposit a check and move funds from one account to another, with nary a coin exchanged. Next, while crossing campus, I negotiated traffic lights, traffic signs, and hand signals from Officer O'Malley. Graduating seniors in regalia, having just received their caps and gowns, were posing in front of every Seahawk sculpture on the university grounds, hamming it up for selfies with friends. As I parked and approached my office, the chime of the clock tower quickened my pace to a meeting with one of my graduate students, to review her thesis progress.

This short vignette relates a sequence of ordinary events—ones that flow so automatically they'd be unlikely to receive a second thought, all seemingly unremarkable. On closer inspection, however, an observant reader might note that every detail of this anecdote involves a symbol. Ranging from the mundane to the sublime (e.g., from braking to love), these symbols were responsible for my very ability to have an ordinary morning. Indeed, much of the richness of our lives, both the precision (e.g., account balances) and the breadth (e.g., abstractions like peace) of our experience, is made possible by the symbolic function of the events that surround us. Imagine the limitations that would necessarily characterize the world of individuals for whom symbolic function is not so easily acquired and, conversely, the benefits that accrue when multiple symbolic functions are generated rapidly. What sorts of steps might be taken to establish symbols when they are lacking, or to maximize their number efficiently? How can a single symbol come to hold a range of different meanings across individuals (e.g., from security to oppression)? And while we may feel that "we know one when we see one," what, exactly, do we mean when we speak of a symbol anyway? The few examples given here should help to illustrate that if we are ever to achieve a healthy and happy world, or even a minimally functional one, it will be important to find answers to such questions. In point of fact, these are the questions that lie at the heart of the study of stimulus equivalence.

Behavior analytic research on stimulus equivalence explores how a person's experience with the environment establishes or changes the function of symbols—that is, their ability to impact our behavior, or their meaning. In doing so, this research helps us understand a great deal about human behavior and about how instruction can be designed and delivered to maximize learning outcomes. Indeed, one of the most significant characteristics of equivalence-based instruction (or EBI) is its generativity. When designed properly, a small amount of explicit instruction can yield an extensive and predictable network of new relations that were not directly targeted by the teaching, which is certainly a win-win scenario for student and teacher alike. The tremendous "bang for the buck" nature of EBI makes it especially relevant to today's pressing need for efficient and effective educational strategies. This chapter further develops the basic concept of stimulus equivalence by illustrating a wide range of equivalence outcomes in applied research

and by outlining key features of the equivalence procedures that practitioners or teachers can incorporate into their own instruction efforts.

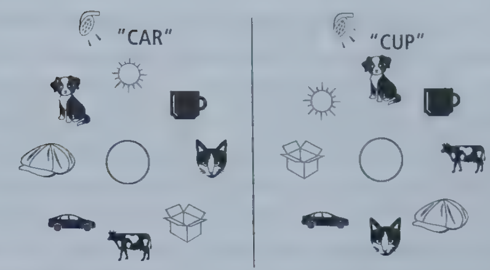


Figure 19.1 An illustration of two trials of the AB matching-to-sample procedure, with the spoken sample stimulus indicated at the top of each diagram.

From "Equivalence Relations and Behavior: An Introductory Tutorial," by M. Sidman, 2009, *The Analysis of Verbal Behavior*, *25*, p. 8. Used by permission.

RESEARCH FOUNDATIONS AND CORE CONCEPTS

The behavior analytic approach to symbolic function known as **stimulus equivalence** made its debut in the early 1970s (Sidman, 1971; Sidman & Cresson, 1973) and received its first full operational treatment in 1982 (Sidman & Tailby, 1982; Sidman et al., 1982). That's not so very long ago in science years. Since that time, probably no other single topic has occupied more attention across the domains of basic experimental, conceptual, and applied behavior analytic science. Let's begin with a brief review of the equivalence phenomenon, before turning our attention to the reasons for its emphasis within the field.

Pioneering Research on Symbolic Stimulus Relations

What Was Taught

Sidman's initial studies on reading acquisition (Sidman, 1971; Sidman & Cresson, 1973) provide a perfect context for an introduction to the basic terms and concepts of an equivalence approach. Because these terms and concepts will play a critical role in the descriptions to follow throughout this chapter, we will spend some time here developing each of them with respect to Sidman's landmark studies. The participants in these experiments were three institutionalized young men (ages 17 to 19) with severe intellectual disabilities. Despite extensive teaching histories, none of the participants had developed any reading skills whatsoever. Sidman targeted this critical functional repertoire with a teaching arrangement known as the **matching-to-sample procedure**, an arrangement that has much in common with everyday classroom teaching approaches. Matching-to-sample is a *discrete-trial procedure*; in Sidman's case, each teaching trial began with the repeated presentation of 1 of 20 different spoken words, referred to as *sample stimuli*, and the illumination of a round key in the center of the teaching apparatus. The participant was required to touch the illuminated key, to indicate that the spoken word had been noticed, a response referred to as an *observing response*. The observing response resulted in presentation of eight simple line drawings, known as *comparison stimuli*, arranged in a circular configuration around the sample key. One of the comparison drawings corresponded to the spoken word presented as the sample stimulus on that trial; for example, given the spoken word "car" as the sample, the comparisons might include drawings of a car, a sun, a cup, a cat, a box, a cow, a cap, and a dog. (See Figure 19.1 for an illustration of two trial presentations.) The participant's job at this point was to select one of the comparison stimuli, again by touching it. Selecting the drawing that corresponded to the spoken word (e.g., selecting the picture of a car after hearing "car") produced pennies and a chime; that is, the corresponding comparison was *discriminative* for reinforcement. Selecting any of the nonmatching comparison stimuli had no programmed consequences. In either case, a brief interval followed, after which the next trial began.

Across trials, each of the 20 words was presented as a sample stimulus. Thus, the comparison stimulus designated as correct, or discriminative, on any given trial varied as a function of the specific sample stimulus presented on that trial. After hearing "car," selecting the picture of a car produced reinforcers; after hearing "cup," selecting the picture of a cup produced reinforcers; and so on. The participant could not earn reinforcers consistently by selecting the correct comparison from the previous trial; correct selections had to "match the sample." Said a bit more technically, stimulus control by the sample stimulus determined the discriminative comparison stimulus. The matching-to-sample procedure thus arranges a *four-term contingency*, the positive elements of which include two antecedent stimuli (the sample stimulus and the correct comparison stimulus), the selection response, and the reinforcer. The four-term contingency brings the more familiar three-term contingency under another level of antecedent stimulus control in that the sample stimulus determines the appropriate three-term, or discriminative, unit. The performance produced by this four-term contingency (i.e., selecting the appropriate comparison on each trial) is known as **conditional discrimination** because discrimination between the comparison stimuli is conditional upon, or depends upon, the sample stimulus present on each trial. It follows that the sample stimuli are also known as *conditional stimuli* in this arrangement. In Sidman's first training phase, then, the discriminative function of each picture was conditional on the word spoken. Also with respect to terminology, Sidman's teaching arrangement could be described as an *arbitrary* or sometimes even a *symbolic* matching-to-sample procedure, because the correct comparison stimulus on a given trial did not depend on any physical relation to the sample (as opposed to *identity matching-to-sample*, where the sample and correct comparison are physically identical, or *oddity matching-to-sample*, in which the correct comparison is the only one physically different from the sample).

It is common in the literature on equivalence to find alphanumeric labels used to specify individual training stimuli. This practice is useful in providing a standard descriptive system that readily accommodates the many variations in stimulus types and

training procedures that are represented in the literature. It can take a little getting used to, but gaining some familiarity with this system will facilitate descriptions of the studies presented throughout the remainder of this chapter and make them easier to follow as a reader. Each alphanumeric label designates a physically unique stimulus, where the letter denotes a stimulus set (e.g., a set of sample stimuli, or a set of comparisons) and the number denotes a potential equivalence class. For example, the spoken-word sample stimuli in Sidman's reading studies could be designated as A1, A2, A3, . . . A20; and across training trials, each sample would be presented along with an array of eight drawings from the B stimulus set, designated as B1, B2, B3 . . . B20. As reviewed previously, during training, selecting comparison B1 in the presence of sample A1, selecting B2 in the presence of A2, and so on, would produce a consequence determined to serve as a reinforcer, while selecting any of the alternative comparisons would result in the intertrial interval only. Once selection of the appropriate B comparison stimulus is demonstrated reliably in the presence of each A sample, the performance is described as an *AB conditional discrimination.* In Sidman's studies, correctly matching pictures to spoken words constituted the AB conditional discrimination.

Sidman's second teaching phase began for each participant once all AB performances were strong. (In fact, his participants were able to match a number of the pictures to words even before training began.) Teaching Phase 2 was conducted exactly as Phase 1 had been, but with one important difference. Instead of the pictures, 20 written words now served as comparison stimuli, with 8 of the words presented on any given trial. Spoken words served again as sample stimuli. (See Figure 19.2.) Thus, Phase 2 was designed to establish an *AC conditional discrimination.*

This training phase took some time to complete, and for two participants (from Sidman & Cresson, 1973) the words were taught in increasing set sizes; first, 9 words were taught to a mastery criterion, then 5 more words were added for a total of 14, and then all 20 words were presented together. When all AB and AC performances were strong, the important questions could then be addressed. Had these young men acquired skills that might reasonably be described as true reading comprehension? Did they really understand the words as symbols, or had they simply learned the rote pairings that had been directly targeted by the training procedures?

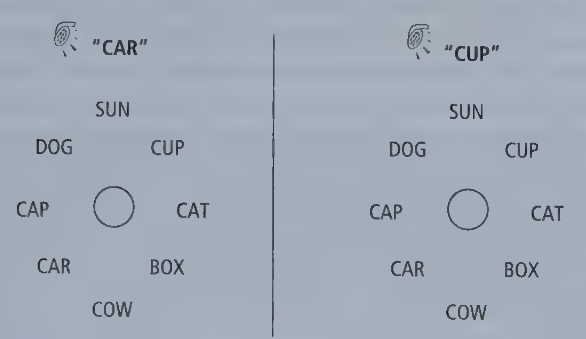


Figure 19.2 An illustration of two trials of the AC matching-to-sample procedure, with the spoken sample stimulus indicated at the top of each diagram.

Adapted from "Equivalence Relations and Behavior: An Introductory Tutorial," by M. Sidman, 2009, *The Analysis of Verbal Behavior, 25*, p. 7.

What Was Learned

To distinguish between the two possibilities, Sidman designed an ingenious series of test or *probe trials*, still presented in the matching-to-sample format, but in the absence of any programmed consequences. In other words, each probe trial was presented under extinction conditions, which allowed interpretations in terms of symbolic, or true matching, function. (In fact, these same tests had been presented at the very start of the studies, prior to any teaching, to determine baseline levels of performance; with the exception of some correct selections on AB trials, as noted above, and some correct picture-naming responses, baseline accuracy levels were uniformly low.) The probe trials presented new combinations of sample and comparison stimuli that were never encountered during training. Thus, any consistent patterns of responding observed on the probes (in the absence of reinforcement) could be described as **emergent stimulus relations**. The term **derived stimulus relations** is sometimes used as a synonym for these emergent performances.[1]

There are four critical types of probe trials in equivalence studies. To illustrate Sidman's training and testing design, Figure 19.3 presents a schematic of the trial types involved, simplified for the sake of space to show only 3 of the 20 stimuli from each training set. Figure 19.4 presents exactly the same information using the alphanumeric stimulus-designation system, to illustrate the relation between the two descriptive approaches. In both formats, arrows point from the sample stimuli, which would be presented individually, to the comparison stimulus array. Solid arrows indicate stimulus relations for which explicit reinforcement contingencies had been arranged during training phases (i.e., the AB and AC conditional discriminations described above). Broken arrows indicate the untrained stimulus relations that could be tested by probe trials. In his 1982 papers (Sidman & Tailby, 1982; Sidman et al., 1982), Sidman provided descriptors for each probe-trial type, taken from the mathematical set-theory definition of equivalence. He reasoned that these behavioral tests captured operationally the abstract mathematical properties said to hold when a set of events are defined as equivalent to one another.

To explain further, after teaching AB and AC conditional discriminations, stimulus **reflexivity** is tested by probes for *generalized identity matching* between each of the experimental stimuli. For example, with stimulus B2 as sample, and stimuli B1, B2, and B3 as comparisons, selecting stimulus B2 would be indicative of reflexivity. Consistent reflexive selections with all of the experimental stimuli (e.g., the pictures and the written words, in Sidman's studies), in the absence of reinforcement for such selections, indicate that each stimulus is functioning as equal to itself. The mathematical statement of reflexivity holds that A = A; thus, generalized identity matching provides an operational definition of the mathematical property. The mathematical statement of **symmetry** holds that if A = B, then B = A or if A = C, then C = A. This property is typically operationalized by probe trials that reverse the roles of the

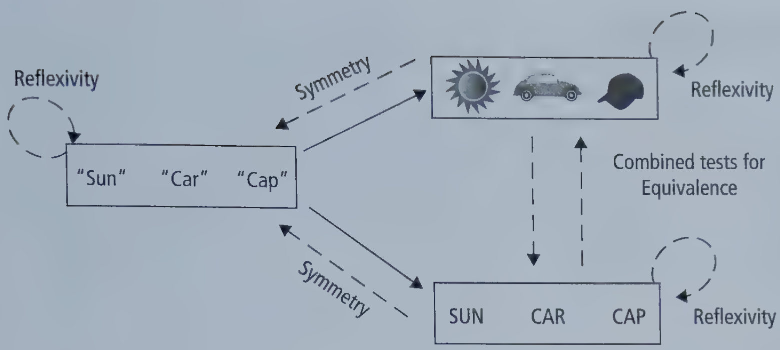


Figure 19.3 Schematic representation of the trained and tested conditional discriminations. Solid arrows indicate trained relations, and dashed arrows indicate tested relations. The arrows point from sample stimuli to comparison arrays.

"Efficiently Establishing Concepts of Inferential Statistics and Hypothesis Decision Making through Contextually Controlled Equivalence Classes" by D.M. Fienup and T.S. Critchfield, 2010, Reproduced with permission of John Wiley & Sons Inc.

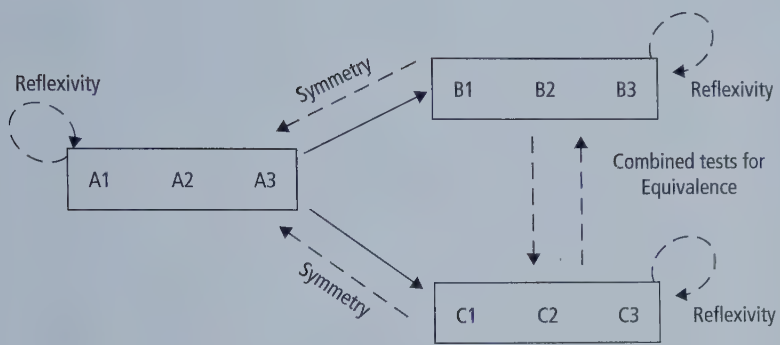


Figure 19.4 Schematic representation of the trained and tested conditional discriminations, with individual stimuli indicated by their alphanumeric designations.

stimuli presented as sample and comparisons during training. For example, after learning to select the appropriate B comparison given an A sample (i.e., selecting B1 given A1, B2 given A2, and B3 given A3), a symmetry probe would present one of the B stimuli as a sample, with A1, A2, and A3 as the comparison choices. Similarly, after establishing selections of the appropriate C comparison stimulus given an A sample, probe trials would present each of the C stimuli as samples with A1, A2, and A3 as the comparison array. Consistent patterns of reversible stimulus selections (e.g., choosing A1 given B1, or choosing A2 given C2), again in the absence of reinforcement for the selections, across all relevant probe trials would demonstrate the property of symmetry. The young men in Sidman's early studies received a modified version of the typical symmetry probes because the A stimuli in their training consisted of spoken words, and it is difficult to present multiple spoken words simultaneously as choices for selection. Instead of the standard symmetry test, then, each of the B and C stimuli was presented individually, and the participant was asked to name it (a tact or textual response, respectively), arguably a more demanding task, as the appropriate response had to be produced rather than selected from an array. Nevertheless, the exciting result was that these boys not only named the pictures correctly but also read the written words; that is, they read the words orally where they had been unable to do so prior to the teaching phases, despite the fact that oral reading had not been directly targeted during training, and despite the fact that no reinforcers were available on the probe trials. This was a striking accomplishment, but did the boys understand the meaning of the words?

A third probe-trial type evaluated this possibility. The mathematical statement of **transitivity** holds that if A = B, and if B = C, then A = C, and when both symmetry and transitivity hold, C = A. This latter arrangement is often described as a *combined test for equivalence*, or even more simply, an **equivalence test**, because both properties are evaluated simultaneously. The participants in the present studies had received AB and AC training; thus, their combined tests for equivalence involved either B stimuli presented as samples with C stimuli as comparisons, or the converse—C stimuli as samples with B stimuli as comparisons (i.e., if A = B, and if A = C, then B = C and C = B). On some probe trials, a picture (from stimulus set B) was presented first and the task was to select the matching written word (from stimulus set C). On other probe trials, a written word was presented first, and the task was to select the matching picture. All 20 of the possible picture-word combinations were evaluated in this way. Consistent and appropriate matching of pictures to words (e.g., selecting a picture of a sun when the sample was the written word "SUN," or selecting the written word "CAR" when the sample was a picture of a car), again in the absence of reinforcement for the selections, across all relevant probe trials would demonstrate the combined properties of symmetry and transitivity. More importantly for these young men, such consistent patterns would also represent true reading comprehension, an indication that the meaning of the words was understood. Sidman (2007) describes observing the first probe trials conducted with these young men; he and his colleagues watched from behind a one-way mirror, all with baited breath. As one correct response after another was demonstrated, the scientists literally danced, and at the end of the session, the research assistant jumped up, hugged the participant, and shouted, "Kent, you can read!" (Sidman, 2007, p. 315). This excitement on the part of the researchers helps to communicate the power and the potential of the outcome that had just been observed. Indeed, all three of the young men involved went on to show the emergent patterns

indicative of oral reading and reading comprehension for the first time in their lives, even though these particular performances had not been directly taught.

To complete our lesson on basic equivalence terminology, when successful performances like those just described are obtained on probe trials of all three types—reflexivity, symmetry, and transitivity (or combined tests for equivalence)—in the absence of reinforcement, the defining criteria for **equivalence-class formation** are satisfied. In the studies at issue here, 20 different stimulus classes, each including a spoken word, a picture, and a printed word (i.e., A1B1C1; A2B2C2; . . . A20B20C20) could be characterized as *equivalence classes* in that the members of each class were perfectly interchangeable within a given context. Said another way, each stimulus from a given class functioned identically, whether as a sample or as a comparison stimulus, in relation to any other member of the same class. It is exactly this emergent interchangeability that allows us to interpret the stimuli within each class as symbolic of one another (e.g., a symbolic relation was demonstrated between the spoken word, "car," the written word, [CAR], and the picture of the car). What's more, this interchangeability was the result of a structured teaching approach involving only a small subset of the total number of relations generated.

Why It Mattered

The reading accomplishments of these three young men thus served to inspire the study of equivalence on multiple fronts. First, new forms of stimulus control were created that had never been targeted directly by the teaching procedures. Consider the productivity of the approach taken in these cases: 20 AB (spoken word–picture) and 20 AC (spoken word–written word) relations either were already demonstrated by these participants or were explicitly established by the reinforcement contingencies of the matching-to-sample training procedure. The resulting performances included those 40 relations, as well as the potential for 60 new reflexivity relations, 40 new symmetry relations, and 40 new combined equivalence relations, for a grand total of 180 discriminated operant units! This production of reliable and predictable patterns of novel stimulus-control relations, particularly when the novel control is not based on physical similarities across the stimuli (as in primary stimulus generalization), is a striking behavioral phenomenon, demanding of research in its own right, just to understand the processes that might be responsible. Prior to Sidman's work (Sidman, 1971; Sidman & Cresson, 1973; Sidman & Tailby, 1982), there was nothing within an extensive behavior analysis literature to suggest such generative capacity after basic four-term contingency training. A tremendous research effort has followed.

A second reason for immediate excitement about Sidman's early findings followed unequivocally from the first, and is perhaps most relevant to the content of this textbook. Given the focus on reading in Sidman's studies, the potential for utilizing an equivalence approach to promote teaching efficiency was immediately apparent. Sidman's early successes had captured that holy grail of applied outcomes—establishing extensive, varied, and flexible repertoires with only a modicum of direct training. The way had been paved for creative extensions of the equivalence approach to behavioral targets across a broad spectrum of academic and functional life skills, a number of which will be reviewed in the pages that follow.

A third striking feature of the equivalence outcome just described is the degree to which it captured the essence of what is meant when we speak of a "symbol." The three defining probe-trial types revealed that an arbitrary sound (i.e., a spoken word), an abstract set of black-and-white markings (i.e., a written word), and a related picture were each treated identically and interchangeably, just as would be expected of a symbol and its referent. This sort of substitutability has often been described in terms of a symbol "standing for," or "meaning the same as," its referent; however, a clear specification of the behaviors necessary for identifying a true symbol had long been lacking (e.g., Sidman, 1994; Wilkinson & McIlvane, 2001; Box 19.1 provides an example of the confusions that often result). Sidman's equivalence criteria had thus established a compelling operational definition of symbolic function. In so doing, these criteria also provided a basis for empirical behavioral analysis of other complex performances critical to much of human behavior (e.g., categorization, semantic correspondence, syntax), as well as a possible approach to the development of (or change to) these sorts of performances as clinical targets. One case in point involves the study of behavioral novelty or creativity. Indeed, creativity has sometimes been erroneously characterized as unamenable to behavioral accounts, given that a history of reinforcement is lacking, by definition, when a behavior occurs for the first time. With equivalence procedures, however, novel behavior is illustrated reliably and in predictable patterns by the emergent performances on probe trials. Other examples of extensions to complex human functioning include equivalence-based models of instruction following, social stereotypes, false memories, and clinical issues such as the spread and treatment of generalized anxiety or phobias, just to mention a few.[2] An important shared characteristic of these complex patterns of responding is the involvement of emergent symbolic function.

Continued Development of Equivalence Research: Key Outcomes Obtained with Equivalence Approaches

Recent years have seen a repeated emphasis on the importance of *translational research*, both for science in general (e.g., Woolf, 2008; Zerhouni, 2005) and for behavior analysis in particular (e.g., Critchfield, 2011; Mace & Critchfield, 2010). A traditional conceptualization of translational research has focused on moving the findings from our basic laboratory science into the hands of practitioners, where our increased understanding can be put to important functional ends. In fact, this pragmatic strategy has been a defining characteristic of behavior analysis since its early formulations. "What we call the scientific conception of a thing is not passive contemplation. When we have discovered the laws which govern a part of the world about us, we are then ready to deal effectively with that part of the world" (Skinner, 1953, p 13). At the same time, behavior analysts have long realized that when those behavioral principles emerging from our labs are applied to significant social problems (Bear, Wolf, & Risley, 1968, 1987), additional critical questions about the principles and their operation inevitably arise. This often necessitates a return

BOX 19.1

But I Know My Dog Understands Me

We often hear heartwarming stories in the media about clever pets—especially dogs—who understand their owner's words so well that they perform life-saving rescues (remember Timmy telling Lassie to go get help?), even calling 911 and barking into the phone. Indeed, well-trained dogs help owners with mental health issues cope with emotional overload; they follow commands from owners with physical disabilities to retrieve dropped items, open and close doors, and provide support for standing and walking; and they fetch medications, phones, and other items for the elderly in emergency situations. There are reports of dogs who recognize more than 1000 human words, and one recent functional magnetic resonance imaging (fMRI) study (Andics et al., 2016) found different types of brain activity when dogs were presented with meaningful words in either a praising or a neutral tone compared to meaningless words presented as praise. Further, many dog owners are convinced that their pets understand not only their words but their emotions as well; we all know pets who respond very differently in the face of an owner's laughter versus tears. Surely these reports, and scads of others like them, document that dogs understand what we say, right?

This question gets right to the heart of what it means to speak of "understanding a word." To pose the issue a bit differently, is there any reason to believe that a dog's response to "Sit," for example, differs from a child's, or from our own, for that matter? In all three cases, the appropriate bodily position may be executed perfectly and immediately upon hearing the verbal command, so observation of that particular interaction does little to reveal a distinction between the three instances. Nevertheless, should we feel comfortable concluding that the dog in question actually understands the meaning of the word "Sit" in the same sense that the child does? Is "Sit" a true symbol for a certain action or posture, or is the dog's "understanding" limited to performing a single response that has been explicitly trained in the presence of a specific discriminative stimulus? These are essentially the same questions Sidman was grappling with in his early equivalence studies (Sidman, 1971; Sidman & Cresson, 1973; described above), when he sought to find out whether the match-to-sample teaching procedure had produced true reading comprehension for the boys who participated. When the boys came to choose the appropriate written word upon hearing it spoken after repeated reinforcement for doing exactly that, did they understand the meaning of the words, or were they simply performing a trained rote response? As with the dog example, accurate selection of the directly trained option, in and of itself, did not reveal an answer. Further indicators, beyond the explicitly reinforced behavior patterns, are needed to make judgments about understanding, comprehension, and symbolic function. To return to the "Sit" example, after learning to take a seat when asked to, a child might then also request that her mother "sit" in an appropriate circumstance, accurately label or describe other people as sitting, match "Sit" with pictures of novel individuals doing so, and point out suitable objects on which one might sit, all in the absence of further teaching. These novel or emergent performances provide the types of evidence necessary to reflect understanding of the word. In contrast, responses to "Sit" for even the cleverest pets would likely be limited to executing the one physical position they had been taught to perform. Dogs may indeed be taught to make specific and even useful responses when presented with many different spoken words, which is impressive, to be sure. It falls short, however, of demonstrating true symbolic understanding.

Of course, operationalizing exactly what is needed to convince us of understanding and comprehension provided the impetus for Sidman's development of equivalence, or symbolic matching-to-sample, as characterized by the properties of reflexivity, symmetry, and transitivity (Sidman & Tailby, 1982). Indeed, ever since those defining features were first introduced, the possibility that equivalence patterns might be demonstrated by (nonhuman) animals has been a topic of considerable research activity. Early work on the topic found little support for responding consistent with the properties of equivalence, particularly symmetry, with species ranging from pigeons (e.g., Hogan & Zentall, 1977), to *Cebus* monkeys (D'Amato, Salmon, Loukas, & Tomie, 1985), to rhesus monkeys and baboons, (Sidman et al., 1982), and even to chimpanzees (e.g., Dugdale & Lowe, 1990). More recent work has focused on the very real possibility that the experimental procedures designed for humans may not be optimal for assessing equivalence-class formation in animals—in other words, we may have been asking questions of our animal subjects in the wrong ways. Methodological arrangements better suited to the strengths and learning styles of a given species (e.g., using olfactory stimuli with rats; Pena, Pitts, & Galizio, 2006) have yielded some promising leads. Interestingly, some of the most compelling evidence to date for equivalence-class formation in animals has come from two sea lions, Rocky and Rio (Kastak, Schusterman, & Kastak, 2001).

The experimental histories of these subjects took place over many years, and included extensive training on a number of related tasks, so the complete set of factors responsible for their strong test performances cannot be specified. You can rest assured, however, that the search for equivalence in animals is alive and well in research labs around the world (e.g., Galizio & Bruce, 2017).

to the lab for further investigation, followed by a new round of refined applications, and so the cycle continues. In this way, translational research represents a reciprocal and dynamic process in which the laboratory and the field inform and inspire each other, leading to increased effectiveness in both domains, and providing for the strength of behavior analysis as a discipline. For present purposes, there is perhaps no better example of a translational research agenda than the one that has characterized stimulus

equivalence. The equivalence-research story is especially notable in this context in that the core phenomenon was brought to light initially in an application, reading instruction, which then demanded laboratory follow-up, rather than the other way around—a valuable exemplar for applied behavior analysts and basic researchers alike. Since Sidman's landmark studies, many hundreds of basic laboratory analyses have investigated the conditions necessary and sufficient for equivalence-class formation, while equivalence applications have benefited learners ranging from young typically and atypically developing children, to college students, to adults with brain damage and dementia. As this body of work developed and progressed, several key outcomes were identified that added fuel to excitement about the scope and power of equivalence approaches for applied behavior analysis. These outcomes include class formation, expansion and merger, transfer of function, and contextual control.

Class Formation

The outcome of primary importance in any study of equivalence is demonstration of the emergent performances that indicate stimulus interchangeability and thus the formation of equivalence classes. The defining criteria for interpretations of equivalence have traditionally included strong and stable class-consistent responding on reflexivity, symmetry, and transitivity/equivalence probe trials in the absence of reinforcement (e.g., Sidman & Tailby, 1982; see the dashed lines in Figures 19.3 and 19.4 for illustration of these critical tests). The most common measure of class formation has been the percentage of probe trials on which a comparison selection is consistent with equivalence, for each defining property. For example, after training AB and AC conditional discriminations, stimulus C1 would be presented as the sample on some probe trials; selecting C1, A1, or B1 would be consistent with equivalence-class formation on probes for reflexivity, symmetry, or combined tests for equivalence, respectively. (Again, see Figure 19.4.)

To illustrate, Figure 19.5 provides data from one of Sidman's original participants (Sidman & Cresson, 1973). The three panels present outcomes on the test types of particular interest in that study: word-name relations (i.e., textual responses, or CA oral naming) in the top panel, with picture–printed word (BC) and printed word–picture (CB) relations (i.e., reading comprehension) in the middle and bottom panels, respectively. The trial types were intermixed in the actual test administration but are separated in this figure so that each performance can be followed individually. Each bar represents the percentage of responses on a given probe-trial type (CA, BC, or CB) that were consistent with the potential equivalence classes, for each of a series of test administrations. (As a reminder, 20 three-member classes were made possible by the training, where each class included a spoken word, a picture, and a written word.)

In this study, four pretest administrations were conducted over the course of approximately 1 year. This type of pretesting is often required in studies of equivalence-based instruction when the stimuli targeted for training are common ones found in everyday environments. Pretest results play an important role in judging whether the performances of interest may have been due to learning that occurred outside the programmed

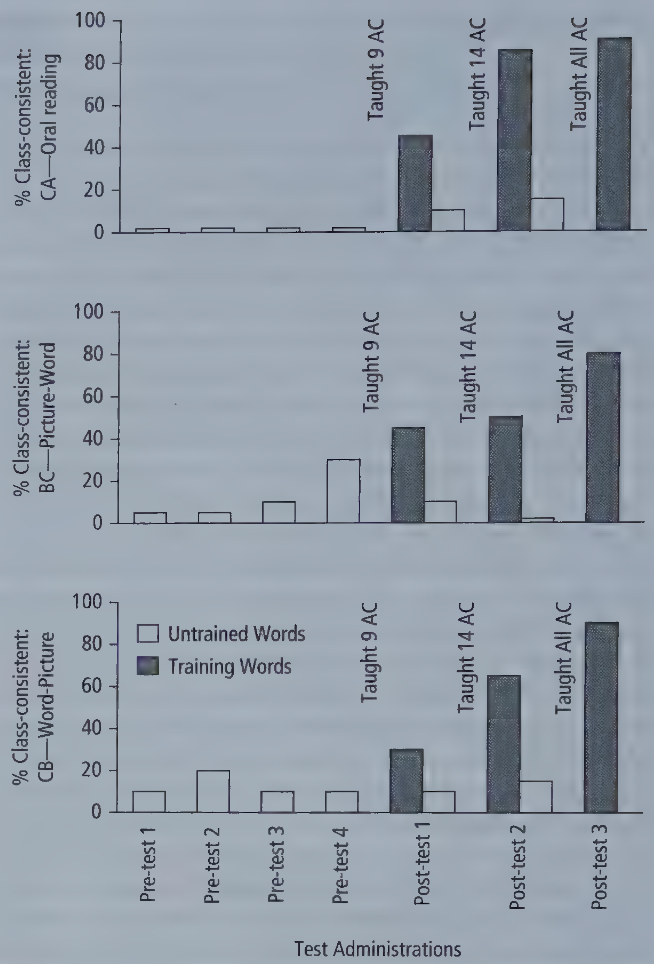


Figure 19.5 Percentage of trials on which class-consistent comparisons were selected for each of three probe types, CA (top panel), BC (middle panel), and CB (bottom panel). Each bar or pair of bars represents data from a single test administration. Open bars show performances on trials with words that had not been involved yet in training; filled bars show performances on trials with words that had been part of training.

Adapted from "Reading and Crossmodal Transfer of Stimulus Equivalences in Severe Retardation," by M. Sidman and O. Cresson, 1973, *American Journal of Mental Deficiency*, *77*, p. 515–523.

instruction under evaluation. In the present case, the percentage of class-consistent responses was consistently low for each of the three probe types across the four pretests, never exceeding chance levels of accuracy, despite repeated practice with each test and despite the fact that the 20 spoken words were being matched with their appropriate pictures (that is, the AB relations had already been learned). Thus, pretest results indicated that the performances of interest were not improving as a result of ongoing extra-experimental experiences for this participant. The AC spoken word–written word relations were taught in three separate instructional phases involving 9, then 14, and then all 20 of the words. A posttest with all 20 words followed each teaching phase. Results from the first two posttests are presented with two bars. The filled bars indicate scores on trials with written words that had been included in the preceding teaching phase, and the open bars indicate scores on trials with written words that had

not yet been taught. By the third posttest, all 20 of the written words had been included in training. Posttest results were similar across the probe-trial types. The percentage of class-consistent responses increased well over pretest levels for each posttest, but only when the trials included a word from the training set. Responding remained at approximately chance levels for those trials involving an untrained word. This important finding indicates that improvements were specific to the instructional procedure, as opposed to some other extraneous factor (e.g., increased motivation or task familiarity, classroom experiences outside the experiment). By the final posttest, class-consistent responding was very strong, at 80% or better, for all probe types, indicating the formation of 20 three-member equivalence classes.

This example illustrates well the type of data and the experimental reasoning that characterize conclusions about equivalence-class formation. While the specific criteria required to make judgments about class formation vary slightly across studies, depending on procedural details, they are always quite high (e.g., 90% class consistency or greater across probe types), and often include further specification that the number of inconsistent responses made on any given probe type is low (e.g., no more than a single error per probe type).

With many participants, equivalence outcomes are observed immediately, on the very first probe tests. However, another relatively common finding, known as *delayed emergence*, is that probe responses become increasingly consistent with class formation as testing continues, sometimes beginning at chance levels and improving to perfect or near-perfect class consistency across test administrations, despite the absence of consequences on these trials (e.g., Sidman, Kirk, & Willson-Morris, 1985; Spradlin, Cotter, & Baxley, 1973). In fact, the data from Figure 19.5 illustrate this pattern. Even for the words that had been included in training, the percentage of class-consistent probe responses increased across the successive posttests. Delayed emergence has important implications for conclusions about equivalence "failures" when probe exposure has been limited, and its frequent occurrence suggests that tests should be repeated when initial performances are weak, especially if the data suggest an improving trend.

Class Expansion and Class Merger

If equivalence-based instruction is to provide a basis for generating truly functional repertoires, class size becomes an issue of interest. In the studies described thus far, three-member equivalence classes (e.g., a spoken word, picture, and written word; Sidman, 1971; Sidman & Cresson, 1973) were produced, this being the minimum number of members necessary to demonstrate all equivalence properties. While clearly valuable, and often a huge improvement from beginning skill levels, three-member classes might not take a child (or adult) very far toward fluency in verbal behavior, numeracy, coin usage, categorization, and so on. Fortunately, equivalence outcomes known as **class expansion** and **class merger** have been studied and documented extensively (e.g., Saunders, Saunders, Kirby, & Spradlin, 1988; Saunders, Wachter, & Spradlin, 1988), providing strong evidence that classes can be increased to include a potentially unlimited number of members. Importantly, as class size increases, the number of generative or emergent performances also grows exponentially.

In class expansion, a new member is added to demonstrated equivalence classes, typically by teaching a new conditional discrimination. One member from each of the original classes serves as either sample or comparison in the new training, and a novel stimulus set serves in the alternative role. To illustrate, consider again the equivalence classes demonstrating Sidman's reading work, where each class included a spoken word, a picture, and a written word (A1B1C1, A2B2C2, etc.). Suppose that for the young men involved, learning the corresponding signs from American Sign Language would also have been useful for their living circumstances. EBI could be used to expand their three-member classes to four members by teaching the young men to match the signs to spoken words. In other words, the signs could be presented as comparison stimuli (e.g., D1, D2, D3) with the spoken words serving as samples (e.g., A1, A2, A3), and an AD conditional discrimination could be taught by reinforcing the appropriate comparison selections. Figure 19.6 illustrates the new training with a solid gray arrow. Class expansion would be evidenced by equivalence-consistent responses (e.g., selecting the appropriate sign given a picture or written word; selecting the appropriate written word or picture given the sign) on probe trials DA, BD, CD, DB, and CB. Figure 19.6 illustrates all of the newly possible emergent relations with dashed gray arrows. Note that after teaching one additional conditional discrimination (AD), a total of six new relations are created for each equivalence class (i.e., AD, DA, BD, DB, CD, and DC), dramatically increasing the impact of a small amount of teaching. Instructing additional conditional discriminations, one by one, can produce a gradual increase in class size from three, to four, to five members or more, even with young children (e.g., Lazar, Davis-Lang, & Sanchez, 1984).

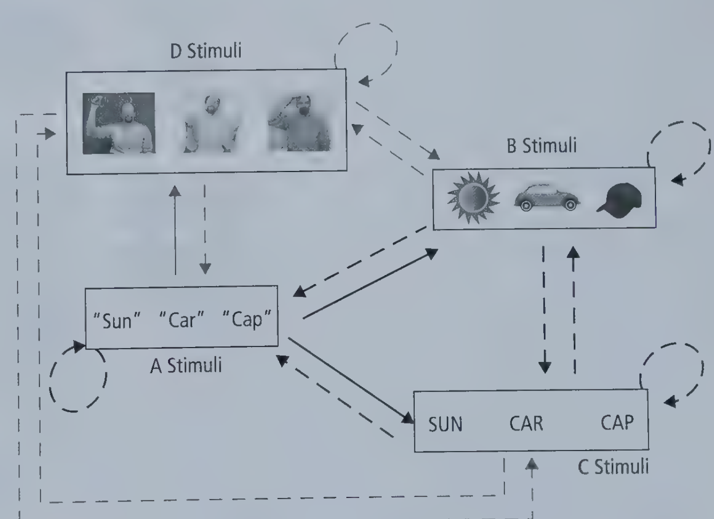


Figure 19.6 Schematic representation of trained and tested conditional discriminations. The solid gray arrow indicates the new conditional discrimination training, and the dashed gray arrows represent the new emergent relations made possible by the training.

In class merger, independent equivalence classes can be combined to produce a larger class by teaching a new but interrelated conditional discrimination. Consider the teaching arrangement illustrated in Figure 19.7, which might be appropriate for young children or individuals with intellectual disabilities who lack coin skills. Here, teaching AB and AC conditional discriminations could result in two three-member equivalence classes, one including the spoken word, "Dime," the printed stimulus, 10¢, and an actual dime (i.e., A1B1C1), and the other including the spoken word, "Quarter," 25¢, and an actual quarter (i.e., A2B2C2).[3] As usual, we would look for evidence of these classes having formed by presenting the test trials represented by the dashed lines. Next, two additional conditional discriminations, DE and DF, are taught. Again, performances on the tests represented by the dashed lines would evaluate the formation of two new equivalence classes, one including the spoken "10 cents," two nickels, and five pennies plus a nickel (i.e., D1E1F1), and the other including "25 cents," two dimes plus a nickel, and five nickels (i.e., D1E2F2). At this point a total of four independent three-member classes have been made possible by the training.

Once class-consistent probe performances confirm the formation of these classes, Figure 19.8 illustrates an additional teaching step that could result in class merger. The bold solid arrow shows that one more conditional discrimination—CE—is taught. Given 10¢ or 25¢ as a sample stimulus, reinforcers would follow selection of two nickels or two dimes plus one nickel, respectively.

Class merger describes an outcome in which stimuli from the originally independent classes A1B1C1 and D1E1F1 are matched to each other, such that a single six-member class (A1B1C1D1E1F1) emerges. In our example, the spoken words, "Dime" and "10 cents," the written stimulus, 10¢, and each appropriate coin arrangement, whether a dime, two nickels, or five pennies plus a nickel, could all prove to be completely interchangeable. Similarly, the six stimuli relevant to the "quarter" class could merge and become interchangeable for a second six-member class (A2B2C2D2E2F2). In short, the four original equivalence classes, composed of three members each, could give rise to two six-member classes. Figure 19.9 represents with dashed gray arrows all of the emergent relations made possible by teaching the CE conditional discrimination, and those that would have to be tested to reach conclusions about class merger. Positive, class-consistent responding across the various probe types would provide not only a basis for important functional skills in coin usage but also exactly the evidence necessary for inferences about understanding or comprehension of the symbolic relations between these stimuli. The network of new performances illustrated also paints a striking picture of the generative potential of equivalence-based instruction, especially as class size increases. Indeed, the balance of solid arrows (directly taught relations) to dashed (emergent relations) for the figure as a whole captures nicely the "bang for your teaching buck" impact of equivalence-based teaching procedures.

Class merger outcomes like those just described have been demonstrated repeatedly in the equivalence literature, and with many populations, including typically developing children and adults (e.g., Sidman et al., 1985; Johnson, Meleshkevich, & Dube, 2014), as well as participants with intellectual disabilities (e.g., Lane & Critchfield, 1998; Saunders, Saunders, Kirby, & Spradlin, 1988). To date, few limits have been identified for the size of equivalence classes that can be generated through expansion or merger, although this would likely depend on the particular participant or participant population. Indeed, it has even been suggested that measures of the number of members that can be added to an equivalence class, either sequentially (as

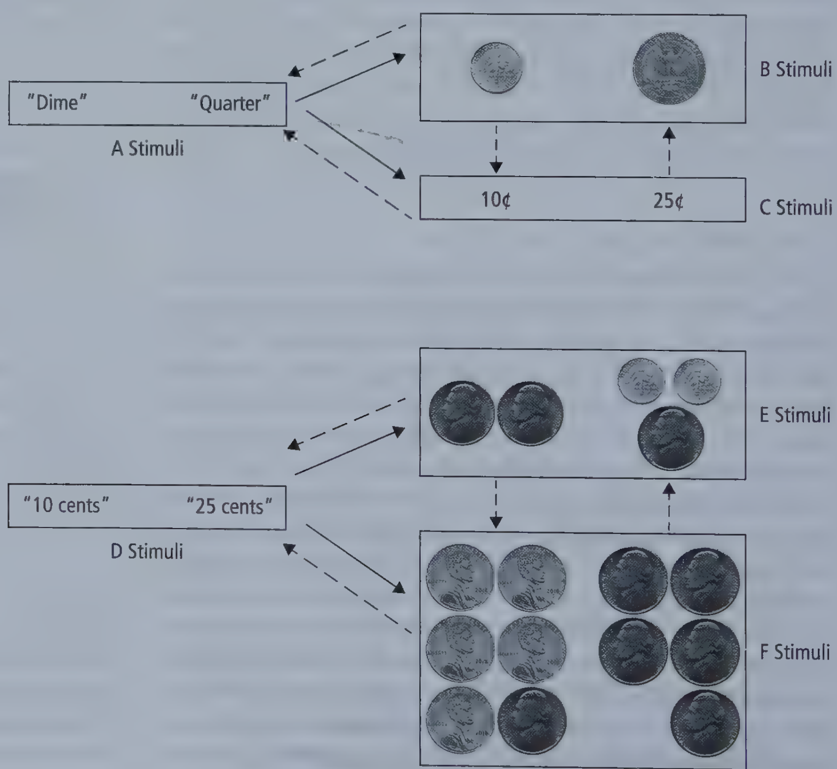


Figure 19.7 Schematic representation of trained (solid arrows) and tested (dashed arrows) conditional discriminations for two training phases, one for AB and AC conditional discriminations and one for DE and DF conditional discriminations.

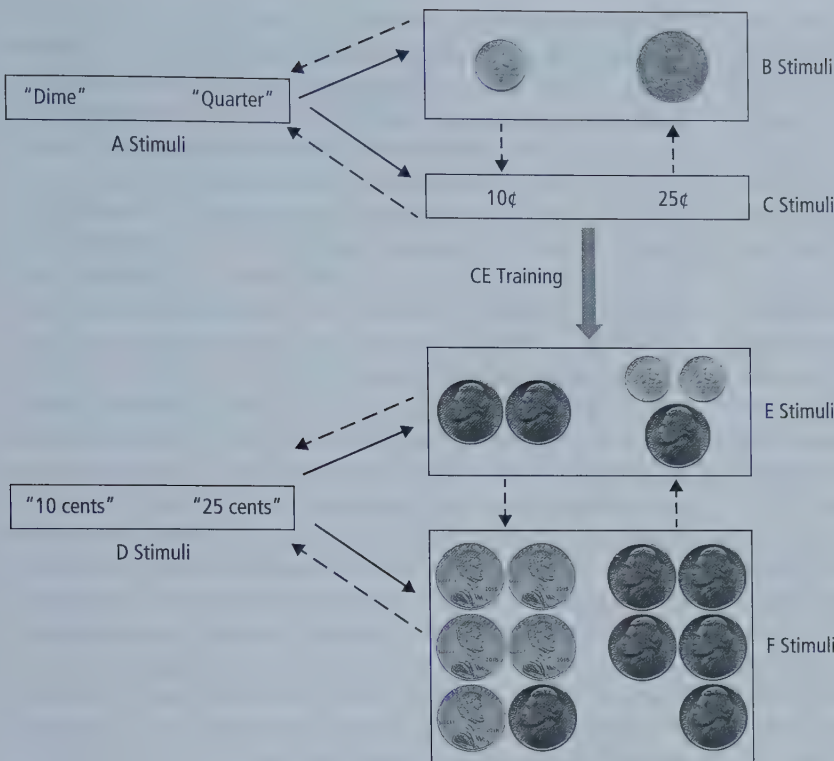


Figure 19.8 Schematic representation of trained (solid arrows) and tested (dashed arrows) conditional discriminations for the first two training phases, as well a third training phase. The bold solid arrow indicates the added CE conditional discrimination training.

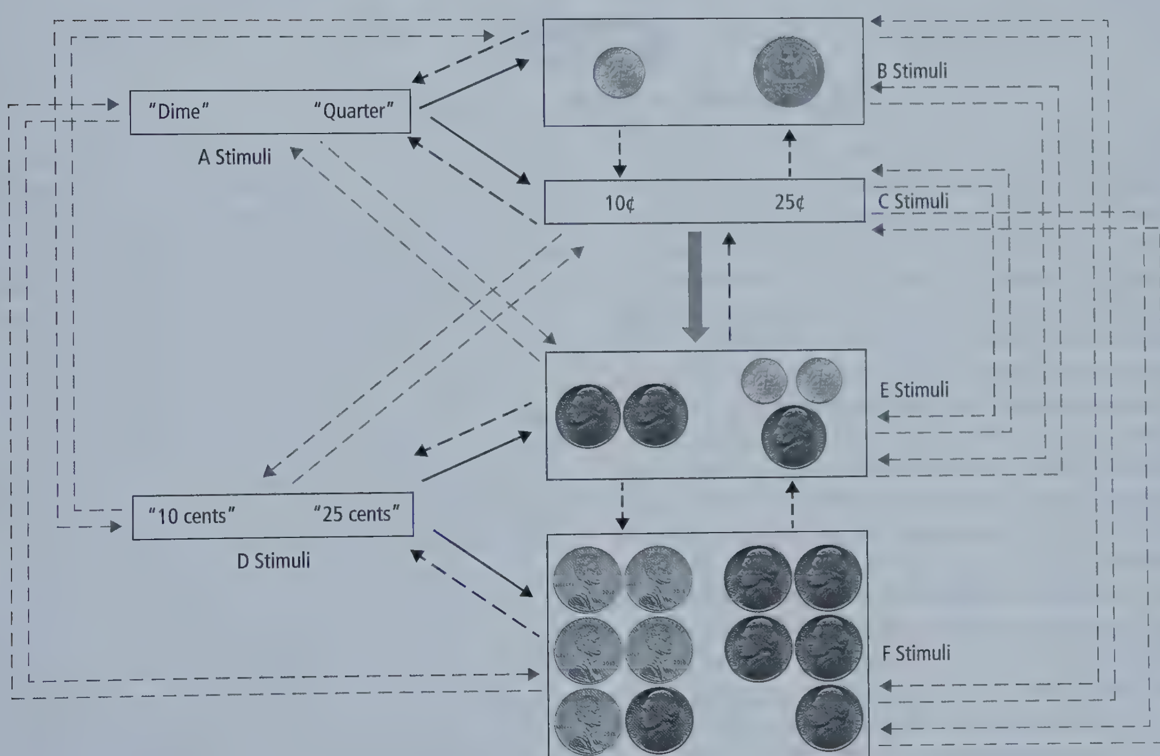


Figure 19.9 Schematic representation of trained (solid arrows) and tested (dashed arrows) conditional discriminations for all training phases. The bold solid arrow indicates the added CE conditional discrimination training that made class merger possible. The dashed gray arrows indicate the many relations that could result from class merger.

in class expansion) or simultaneously (as in class merger), might provide a more valid assessment of cognitive development or of recovery from brain injury than the standardized tests currently in use (e.g., Sidman, 1986a; 1994).

Transfer of Function

A third major finding from research on stimulus equivalence, described as **transfer of function**, adds even further power to the generative potential of EBI. When a group of stimuli are said to be

equivalent, the implication is that each of the stimuli in the group will function in the same manner: that each is interchangeable with any other. As we have reviewed, this interchangeability of function is assessed by tests of reflexivity, symmetry, and transitivity/equivalence in which each member of the class proves to serve equally well in relation to any other member of the class, no matter which stimulus combinations are presented as sample and comparison. But what if a new function—that is, a stimulus-control effect other than that of sample or comparison in a match-to-sample task—was taught for one member of an equivalence class? If the stimuli that compose a given class are truly equivalent, it would follow that the new function should be shared by all members of the class—in other words, that the function should transfer to all class members.

Let's continue with the coin-equivalence example from Figure 19.9 and assume that our students have demonstrated the two possible six-member classes. As a next teaching step, we'll establish a new function for the dime and the quarter (B1 and B2). First, we'll teach that the dime can be traded for two pieces of candy and that the quarter can be traded for five pieces of candy. When this lesson is mastered, we'll provide a choice between the dime and the quarter: "Which one would you like to have?" Selection of the quarter would confirm that our teaching procedure established a new function for each coin—that the quarter is better than or worth more than the dime. Subsequent tests would then present the same choice with each pair of stimuli (A1 and A2, C1 and C2, D1 and D2, etc.). Consistent selections of A2, C2, D2, E2, and F2 (instead of A1, C1, D1, E1, or F1) would demonstrate the transfer-of-function effect in that each member of the quarter class is preferred to, or better than, a member of the dime class, despite the absence of direct "value" training for any of these class members.

The transfer-of-function effect has been demonstrated in a great many ways, shapes, and forms, and for many different stimulus functions (e.g., Green, Sigurdardottir, & Saunders, 1991; Kohlenberg, Hayes, & Hayes, 1991; Smeets & Barnes-Holmes, 2003), highlighting the possibilities for emergent equivalence-based performances that go far beyond the matching-to-sample context. Indeed, transfer of function is arguably among the most important advantages of classes or categories; once a class has formed, anything new learned about one member of the class can hold for all members of the class, without additional training. However, as illustrated in Box 19.2, transfer of function can be implicated in the spread of both adaptive and problematic response patterns. The transfer effect is accounted for from the perspective of equivalence in terms of class merger (e.g., between the established equivalence class members and the class of stimuli that serves a particular function; e.g., Sidman, 1994, 2000). Equivalence-based applications designed to take advantage of transfer-of-function effects are still somewhat limited, but the potential is clear.

BOX 19.2

Are Equivalence Classes Ever a Problem?

A chapter on designing and implementing EBI will necessarily focus on productive uses of this teaching approach. A completely fair question, however, is whether equivalence-class formation might ever have a negative impact on an individual's functioning. As you might guess, in a world in which bad things exist, equivalence classes that include bad things as members are also a possibility. And as we've just reviewed, if one member of an equivalence class has a particular function (e.g., if one member is "bad"), the other members of the class are likely to share that same function—the transfer-of-function effect. As a result, physically dissimilar stimuli might come to be treated as "bad" even if they have never had any direct pairing or correlation with aversive events. Of course, establishing classes of bad things can often be useful, even life-saving. Any parent would be delighted if her child immediately avoided items marked with the signs shown in Figure 19.10, after learning to steer clear of POISON, for example.

In contrast, consider a case in which a college student experiences some difficulty while taking a test in a biology class. For question after question, she realizes that her study the night before was far too superficial, and she cannot compose a single coherent answer. As her heart begins to race and her breathing becomes more rapid, her thinking is increasingly distracted. She rushes from the room with most of the test booklet still incomplete, and thus escapes the unbearable situation. We could all agree that this scenario describes a highly aversive experience. Unfortunately for the student involved, the aversiveness doesn't stop there. She begins to notice the same increased respiration and heart rate with respect to other classes, and not only when taking a test. Preparing for a test, or just the announcement of a quiz, exam, or final, can initiate the cycle of intense emotions. Study patterns are disrupted and absences increase, both for exams and for regular class days (pop-quiz possibilities . . .). Course syllabi with test dates and final exam schedules even become difficult to look at, and registering for a new semester looks bleak.

Patterns like this one—often described in terms of test anxiety—are not uncommon among college students. How might equivalence classes play a role? One possible interpretation (e.g., Augustson, Dougher, & Markam, 2000; Dougher, Augustson, Markham, Wulfert, & Greenway, 1994; Friman, Hayes, & Wilson, 1998) goes as follows: For college students, it is not hard to imagine an equivalence class that includes actual tests; the spoken and written words "Test," "Exam," "Quiz," and

Figure 19.10

"Final"; materials describing schedules of such events; and even larger academic units comprising these events (like courses or semesters). The various members of this suggested class share few physical features but may be interchangeable within a given context. To the extent that they are equivalent, then, a frightening experience with one member of the class could result in strong emotional reactions to other members of the class; that is, the function of one member could transfer to them all.

The feasibility of this conceptual account of anxiety has received some strong support from laboratory models. To illustrate (Dougher et al., 1994), a group of college students demonstrated the formation of two four-member equivalence classes (A1B1C1D1 and A2B2C2D2) following match-to-sample training with abstract stimuli. Trials were then presented with either stimulus B1 or B2 on the computer screen. When B1 appeared, a brief electric shock was also presented; when B2 appeared, no shock occurred. In the next phase of the study, the other class members were presented, one at a time. Stimuli C1 and D1 elicited skin conductance responses (a common laboratory measure of emotional reaction) similar to those produced by B1, while C2 and D2 did not (Dougher et al., 1994), demonstrating the transfer of an elicitation function from B1 to C1 and D1. In a second experiment, the eliciting function of B1 was extinguished by presenting that stimulus multiple times in the absence of shock. Subsequent tests with each stimulus presented individually showed that the eliciting function of C1 and D1 had also extinguished. These interesting results were interpreted as a possible model of phobic response patterns, in which relatively innocuous settings or events can come to occasion severe, even debilitating, anxiety and fear, despite the sufferer reporting no bad experiences with the setting or event at issue. The study confirms that a negative experience need not have taken place in the presence of a current anxiety trigger, depending on the stimuli with which it shares equivalence-class membership. That extinction functions could also transfer across class members is potentially relevant to the effectiveness of certain clinical treatments for phobias (Dougher et al., 1994), such as systematic desensitization (e.g., Koegel, Openden, & Koegel, 2004; Wolpe, 1961).

In sum, that same generativity of equivalence classes that provides for great effectiveness in creating productive repertoires has the potential to result in an equally broad impact of unwanted learning experiences. A similar analysis has been applied to social stereotypes—both positive and negative—where learning about a characteristic of one member of an established group is then extended to other individuals on the basis of nothing more than their group membership (e.g., Kohlenberg et al., 1991; Moxon, Keenan, & Hine, 1993).

Contextual Control

An essential characteristic of many everyday classes or categories is that class composition can shift, sometimes dramatically, depending on the context. For example, the events that might be included in a class of "good plays" would be altered considerably if conversation changed from sports to theater. Similarly, a list of musicians would be sorted quite differently, depending on whether one was asked about musical genre, time period, or nationality. In other words, the very same stimuli can be members of different classes, depending on the context. And, of course, we often note that the meaning or function of a word or phrase depends on the context in which it occurs: We respond differently to the word "right" when we follow directions to a destination and when we receive feedback on an answer to an exam question. If the study of equivalence classes is to be useful in our understanding of naturally occurring classes and categories, including language, this same sort of flexibility in class membership must be captured—and it has been, in investigations of **contextual control** over equivalence classes (e.g., Bush, Sidman, & de Rose, 1989; Lynch & Green, 1991; Wulfert, Greenway, & Dougher, 1994). Box 19.3 describes the role of contextual control in some common (pun intended) forms of humor.

Contextual control in these examples is the result of a more complex teaching arrangement known as a five-term contingency (e.g., Sidman, 1986b, 2000), where the match-to-sample performances described previously are brought under the control of an additional antecedent stimulus. Following from our musicians example, with a context of "Nationality" and the Rolling Stones as a sample (A1), Fairport Convention (B1) would be the appropriate choice (both bands are British); with Bob Dylan as a sample (A2), The Doors (B2) would be the best match (as American musicians). Conversely, if the conversation topic (i.e., the context) turned to music genre, The Doors (B2) would be matched to the Rolling Stones (A1; both rock), while Fairport Convention (B1) would be matched to Bob Dylan (A2; as folkies).

Said a bit more technically, selecting a given comparison stimulus (e.g., B1) is reinforced in the presence of a particular sample (e.g., A1) in one condition (i.e., Context 1), while selecting a different comparison stimulus (e.g., B2) in the presence of the same sample (e.g., A1) is reinforced in a different condition (i.e., Context 2). Similarly, selecting B2 is reinforced given sample A2 in Context 1, while selecting B1 is reinforced in Context 2. The correct match thus depends not only on the sample stimulus presented but also on the context. The question of interest for present purposes is whether training conditional discriminations under contextual control will give rise to equivalence classes that are also under contextual control.

To date, work with contextual control over equivalence classes has primarily involved typical adult participants, so extending this research to other participant populations will be important. However, there has been some use of contextual control procedures in equivalence-based instructional packages designed for college coursework. For example, Fienup and Critchfield (2010; see also Fienup, Critchfield, & Covey, 2009) taught concepts of inferential statistics and hypothesis testing to college students using a computerized equivalence approach, with tests on the lesson content presented before and after training. Figure 19.11 illustrates the teaching and testing arrangement, with trained relations shown

BOX 19.3

Why Are Puns Funny?

A pun is defined as "the usually humorous use of a word in such a way as to suggest two or more of its meanings or the meaning of another word similar in sound" (Merriam-Webster). Sometimes puns make us laugh, other times they make us groan; we may even apologize for a pun (e.g., as in "no pun intended"), as if punning is in bad taste. It has been argued variously that "Puns are the highest form of literature" (attributed to Alfred Hitchcock) and that "Puns are the last refuge of the witless" (attributed to Samuel Johnson), but authors from Shakespeare to Lewis Carroll to James Joyce to Oscar Wilde have all been recognized for their puns. Why would this form of word play evoke such different reactions, and what is it really that tickles us, or not, when puns are used?

Consider these examples:

- I wondered why the baseball was getting bigger. Then it hit me.
- I saw an ad for burial plots, and thought to myself, this is the last thing I need.
- "Denial ain't just a river in Egypt." (Mark Twain)
- A Freudian slip is when you say one thing but mean your mother.
- You matter. Unless you multiply yourself by the speed of light. Then you energy.
- "I went to a place to eat. It said 'breakfast at any time.' So I ordered French toast during the Renaissance." (Stephen Wright)
- And my personal favorite—"Time flies like an arrow. Fruit flies like a banana." (Groucho Marx)

Are we groaning yet? In everyday terms, we could describe each of the instances above as including at least one word with multiple meanings, either for the same word (e.g., "matter" can refer to physical substance or to importance, "flies" can refer to the insect or to speed) or for one that sounds the same (e.g., "denial" and "the Nile," or "your mother" and "another"). In more technical terms, we could say that, in each case, the word or words on which the joke depends are members of more than one equivalence class. For example, the spoken and written word "flies" is a member of one equivalence class that includes nouns such as the Latin name Diptera; a wide range of winged insects with large compound eyes like house flies, horse flies, and tsetse flies; and descriptors like "pollinators" or "pests." It is also a member of an equivalence class that includes verbs that specify moving through the air, moving at a high speed, moving freely, or simply passing quickly. Typically, when a given word has more than one meaning (as most do), the context in which the word occurs determines its function, and we have seen how contextual control can be responsible for determining the class membership of a particular stimulus (i.e., in this case, whether "flies" is responded to as a member of its noun class or its verb class).

The humor of puns, then, can come about in at least two ways. First, either an ambiguous context is provided such that more than one class membership is possible, or the context is explicitly appropriate to more than one equivalence class; the first two examples above illustrate this type of pun. Alternatively, the pun can be deliberately set up to evoke conflicting classes, as in the last three examples (e.g., Groucho first sets the stage for the verb class of "flies," and then switches to the noun class without much additional warning or context). Where skillful writers or orators typically use context to make the impact of their words more precise, puns from this latter category manipulate the reader/listener's response in opposing directions, thereby creating a smile, but perhaps also the cha-grin of those who are not fans of puns. (Sorry.)

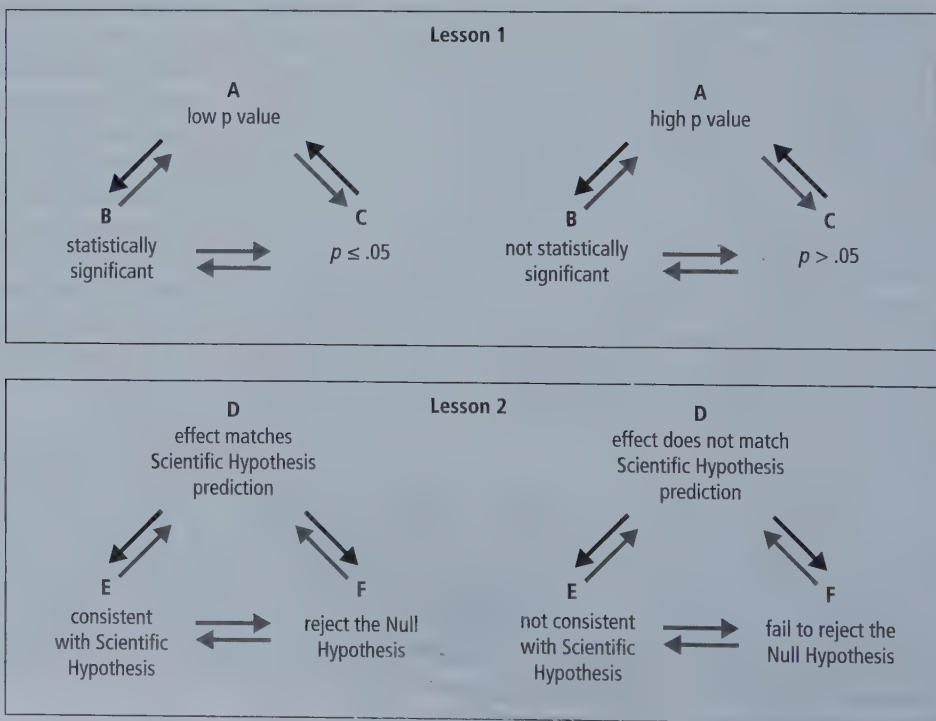


Figure 19.11 Schematic representation of trained and tested conditional discriminations from Lessons 1 and 2, with trained relations indicated by black arrows and tested relations indicated by gray arrows.

"Efficiently Establishing Concepts of Inferential Statistics and Hypothesis Decision Making through Contextually Controlled Equivalence Classes" D.M. Fienup and T.S. Critchfield, 2010, Reproduced with permission of John Wiley & Sons Inc.

by dark arrows and tested relations shown by gray arrows. Lesson 1 established two three-member equivalence classes involving basic concepts related to inferential statistics (A1B1C1 included "low p value," "statistically significant," and "$p < .05$"; and A2B2C2 included "high p value," "not statistically significant," and "$p > .05$"). Lesson 2 did the same with respect to reaching statistical conclusions based on a scientific hypothesis and the direction of an effect in the data (i.e., D1E1F1 and D2E2F2, where D stimuli described a hypothesis and a result, E stimuli were "consistent . . ." or "not consistent with hypothesis," and F stimuli were "reject . . ." or "fail to reject the Null Hypothesis"). Finally, the top panel of Figure 19.12 illustrates Lesson 3, in which contextual control by the A stimuli ("low p value" or "high p value") over DE and DF matching was trained. Note that the correct E or F choice depended on both the D sample and the A contextual stimulus. In other words, selecting "consistent with Scientific Hypothesis" or "reject the Null Hypothesis" was reinforced only in the presence of "effect matches hypothesis prediction" and "low p value". For any other combination of sample and contextual stimulus, selection of "not consistent with Scientific Hypothesis" or "fail to reject the Null Hypothesis" was reinforced. The second and third panels of Figure 19.12 illustrate the test trials used to evaluate the transfer of contextual control to the B and C stimuli.

All 10 students demonstrated the ABC and DEF equivalence classes made possible by Lessons 1 and 2, as well as the relations trained under contextual control in Lesson 3, improving from moderate pretest scores at best to perfect or near-perfect scores in the posttest. Nine of the 10 students also demonstrated high percentages of emergent contextual control by the B and C stimuli. Figure 19.13 presents scores for each student on the two contextual transfer tests (illustrated in Figure 19.11) for all trials (left panels) and for those trials that required control by Lesson 3 (right panels). In this demonstration, then, many sophisticated performances emerged (4.7 times more than the number of relations directly taught) after approximately 45 minutes of training time. What's more, these emergent performances reflected an understanding of the interrelations that define basic statistical reasoning, even in the absence of instructor explanations—an

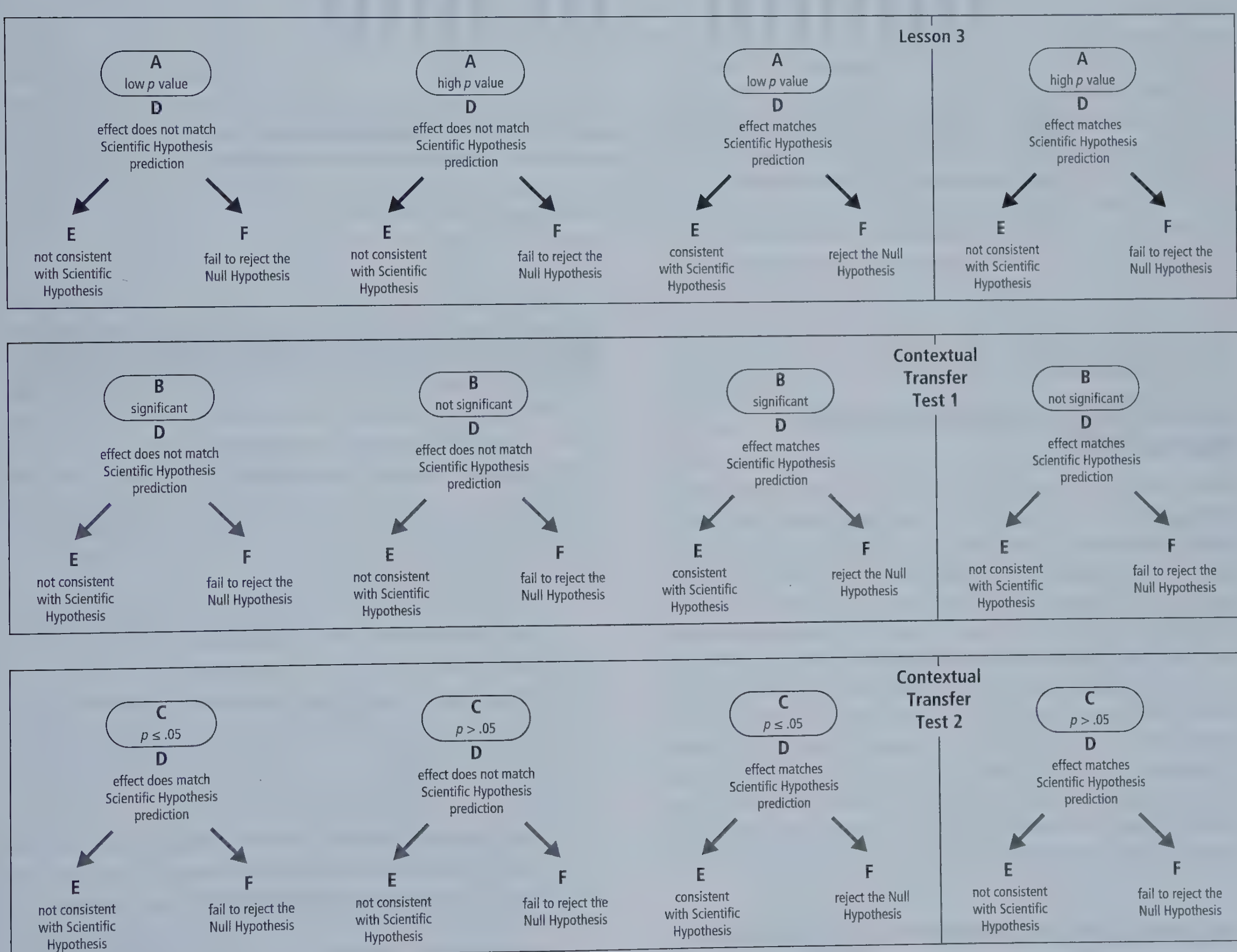


Figure 19.12 Schematic representation of trained contextual control relations from Lesson 3 (top panel) and contextual control tests (middle and bottom panels), with trained relations indicated by black arrows and tested relations indicated by gray arrows. Note that the tested contextual stimuli were never part of contextual control training, but were part of equivalence classes established in Lesson 1.

"Efficiently Establishing Concepts of Inferential Statistics and Hypothesis Decision Making through Contextually Controlled Equivalence Classes" D.M. Fienup and T.S. Critchfield, 2010, Reproduced with permission of John Wiley & Sons Inc.

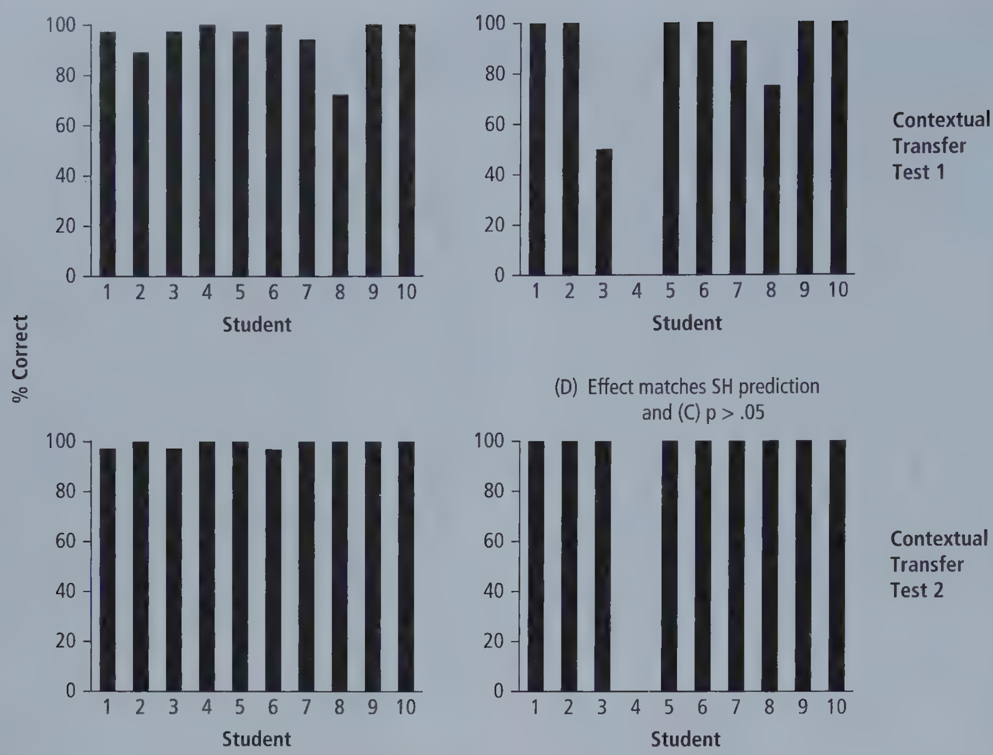


Figure 19.13 Bars represent the percentage of trials on which responses were consistent with emergent contextual control for each student on the two contextual transfer tests (illustrated in Figure 19.11) for all trials (left panels) and for those trials that required control by Lesson 3 (right panels).

"Efficiently Establishing Concepts of Inferential Statistics and Hypothesis Decision Making through Contextually Controlled Equivalence Classes", D.M. Fienup and T.S. Critchfield, 2010, Reproduced with permission of John Wiley & Sons Inc.

impressive illustration of the generativity and nuance made possible with contextual control.

DESIGNING EQUIVALENCE-BASED INSTRUCTION

Equivalence outcomes are impressive, but they do not happen by magic. Research has shown a number of training and testing variables to be important for the successful generation of equivalence classes, whether in the laboratory or in the field. To illustrate these variables, this section of the chapter describes a series of decisions that need to be made in planning effective procedures for EBI, presented in the order in which they would be implemented. In point of fact, the equivalence outcome is quite robust, having been demonstrated with a wide variety of procedural variations, so there is no single route to strong EBI design. A number of problems can be avoided with careful design, however, and the discussion that follows will highlight some of these potential complications. Table 19.1 outlines the major design decisions that guide the planning for EBI training and testing.

TABLE 19.1 Equivalence-based Instruction Design Decisions

I. *Training Considerations*
   A. What stimuli will be used and how many?
   B. How will the stimuli be presented?
   C. Will a sample- or comparison-observing response be required?
   D. What instructions will be given?
   E. How will balanced trial arrangements be ensured?
   F. Are special training steps needed?
   G. What training structure will be used?
   H. What mastery (accuracy) criteria will be used for training?
   I. How will consequences be scheduled during training?

II. *Testing Considerations*
   A. How will the probe trials be presented?
   B. How will consequences be scheduled?
   C. How will the probe-trial types be ordered?
   D. What will be the criteria for equivalence-class formation?

Training Baseline Conditional Discriminations

The most common approach to generating equivalence classes involves teaching a minimum of two interrelated conditional discriminations with the matching-to-sample procedure. Arranging these four-term contingencies requires a number of methodological choices, including the selection and presentation of stimuli, observing responses, instructions, trial-type arrangements and sequences, instructional programming, training structure, mastery (accuracy) criteria, and consequence delivery.

Selecting and Presenting Stimuli

A wide range of stimulus types has been used effectively in equivalence research. In laboratory studies where the emphasis is on experimental control, unfamiliar line drawings (e.g., Dougher et al., 1994), abstract figures or objects (e.g., Pilgrim & Galizio, 1990), letters from a foreign alphabet (e.g., Sidman & Tailby, 1982), or meaningless nonsense syllables are common (e.g., Fields, Hobbie, Adams, & Reeve, 1999), to rule out the possibility that performances demonstrated on probe trials are due to experience with the stimuli outside the lab. Although auditory and visual stimuli are probably the most convenient options, many different stimulus modalities have been represented as equivalence-class members, including tactile stimuli (hidden objects contacted only by touch; e.g., Bush, 1993), odors (e.g., Annett & Leslie, 1995), tastes (e.g., Hayes, Tilley, & Hayes, 1988), and drug effects (DeGrandpre, Bickel, & Higgins, 1992). In applications of EBI, the choice of stimuli will necessarily depend on the nature of the skills being targeted.

Again, the options are virtually limitless. Early language training programs might include objects or pictures, spoken words, and written words, as illustrated in Sidman's early work (Sidman, 1971; Sidman & Cresson, 1973), in either native or non-native languages (e.g., Joyce, Joyce, & Wellington, 1993), while programs designed to establish basic math skills might include quantities of objects, numerals, and spoken number names, or fractions, decimals, and pictorial representations of fractions (e.g., Lynch & Cuvo, 1995). In other illustrations, Braille letters, printed letters, and spoken letter names became equivalence-class members for children with degenerative visual disorders (e.g., Toussaint & Tiger, 2010); the photographs, office name-plates, and spoken names of therapists became equivalent for adults with brain damage (Cowley, Green, & Braunling-McMorrow, 1992); and for college students, the name of brain lobes, specific lobe functions, and brain diagrams became interchangeable class members (Fienup, Covey, & Critchfield, 2010). Applied research with stimulus sets such as these typically incorporate control procedures to rule out extra-experimental learning with the stimuli as responsible for probe performances. Strategies include pretests with each possible stimulus relation, testing all possible stimulus relations before and after teaching each of a series of subsets of the stimuli, and simultaneous multiple-baseline designs.

Another important decision in conducting equivalence work involves the number of stimuli to use. Once again, the best answer here will depend on the participant population and the instructional goals. A number of variations are represented in the literature, but in many cases a reasonable starting point for teaching a first conditional discrimination is to present three different sample stimuli (i.e., A1, A2, and A3) across trials with three different comparison stimuli (i.e., B1, B2, and B3) on each trial. At first glance, beginning with three stimuli for each set may seem counterintuitive—why not start more simply, with just one or two samples and comparisons? The issue is that doing so can allow for the inadvertent reinforcement of stimulus-control patterns that conflict with those being targeted by the training. Consider a case in which sample A1 and comparisons B1 and B2 are presented on every trial for a block of 20 trials, and selection of B1 produces a reinforcer on each trial. A participant might learn quickly to select B1, but doing so would require only a simple discrimination between the two comparisons. Sample A1 has no necessary role in determining the appropriate comparison selection under these conditions, as B1 is always correct. Not only is establishing control by A1 over selection of B1 unlikely here, but control by sample stimuli more generally (e.g., as in control over the participant's attending or observing) could be decreased as well, leaving conditional discrimination more difficult to achieve in subsequent teaching steps.

Next, consider a case in which two sample stimuli (e.g., A1 and A2) alternate irregularly across trials in a block with two comparison stimuli (e.g., B1 and B2) presented on every trial. Now control by the sample is necessary to produce a reinforcer on each trial (that is, selecting B1 is reinforced only in the presence of sample A1, and selecting B2 is reinforced only in the presence of sample A2). In this case, however, it is possible for correct selections to be made, and thus reinforced, in either of two ways. With A1 as the sample, for example, a child could learn to either *select* B1 (known as *Type S* or *select responding*) or to *reject* B2 (known as *Type R* or *Reject responding*) and simply choose whatever else is available. In the first case, the reinforced relation is between sample A1 and comparison B1; in the second case, the reinforced relation is between sample A1 and comparison B2—two different forms of stimulus control (also described as different **stimulus-control topographies**; Dube & McIlvane, 1996; McIlvane & Dube, 2003). The first form of stimulus control would be consistent with the teacher's goal; that is, *stimulus-control topography coherence* would hold between the teacher's target and the child's behavior. If only the second form of stimulus control was established, a lack of coherence would be present, and any resultant equivalence classes would necessarily differ from those targeted by the teacher as well (e.g., A1 and B2 could be members of the same class; see, for example, Carrigan & Sidman, 1992). Using more than two comparison stimuli decreases the likelihood of exclusive Type R control, because the number of reject relations required to produce reinforcement on any given trial increases with each comparison added to the array. For example, given A1 as sample, both comparisons B2 and B3 would need to be rejected from a three-choice array, while only a single Type S relation (e.g., A1B1) is required to produce reinforcers (e.g., Sidman, 1987). In certain circumstances, limiting comparisons to only two or presenting only a single sample stimulus for a series of trials may be appropriate, but the teacher will then need to be watchful for evidence of conflicting stimulus control, and additional training steps may be necessary to ensure a shift in stimulus control.

Given these issues, it follows that presenting at least three comparison stimuli per trial is generally recommended. As a complete demonstration of all three properties of equivalence (i.e., reflexivity, symmetry, and transitivity) requires a minimum of two interrelated conditional discriminations (e.g., AB and AC) with at least three stimulus sets (e.g., A, B, and C stimuli), a total of nine different stimuli would be necessary as a starting point. From there, the number of stimuli per set can be increased as appropriate to the task at hand (e.g., remember that Sidman's first studies used stimulus sets of 20 each, for a

total of 60 stimuli), and as we have seen, the number of sets can be increased as well.

Equivalence-based instruction is also flexible with respect to the mode of stimulus presentation. Modern technology makes computerized matching-to-sample procedures relatively straightforward to design and implement; both the basic laboratory and applied literatures are replete with examples (e.g., see Green & Saunders, 1998; Saunders & Williams, 1998). The sample stimulus is typically presented either at the top of the computer screen or in a center position, with the comparisons presented either in a horizontal row below the sample or in the corners of the screen. The participant's responses (typically mouse clicks, key presses, or touches) are recorded automatically, with auditory and/or visual consequences programmed to follow immediately. Computer presentations have the advantage of ensuring procedural integrity, as all teaching and testing parameters are pre-programmed and consistently applied. Indeed, the promise of developing computerized equivalence-based teaching programs for specific targets in educational settings (e.g., reading, spelling, coin equivalences) has been long noted (e.g., Sidman, 1994), freeing teacher time from basic lessons to more individualized instruction. To say that current application platforms and Internet technologies have significantly enhanced this potential would be an understatement (see, for example, Barron, Leslie, & Smyth, 2018).

Sophisticated computer programming is hardly necessary for equivalence approaches, however, and many elegant examples of "table-top procedures" have been utilized effectively. For example, stimuli have been presented manually in a Wisconsin General Testing Apparatus (Harlow, 1949), where the experimenter and participant sit on opposite sides of a wooden partition. Raising a door at the base of the partition allows a tray holding stimulus objects to be presented to the participant. For example, the sample stimulus might be positioned at the top of the tray, with three comparison stimuli arranged in a row at the bottom of the tray. A concave well below each stimulus object allows a reinforcer (e.g., a small edible, a penny, or a token) to be hidden, so that removing the appropriate comparison provides access to the reinforcer (e.g., Pilgrim, Chambers, & Galizio, 1995). Similar strategies have included a large teddy bear or cloth screen used in place of a wooden partition to block any potential experimenter cuing, with stimulus cards arranged in a systematic format in front of the participant (e.g., Horne, Lowe, & Randle, 2004). In other arrangements, sample and comparison stimuli have been presented as pictures on a page of a three-ring binder or notebook (e.g., with a sample picture at the top of the page and comparison stimuli in a row at the bottom), where each page presents a new trial, or even more simply, by arranging individual objects or cards with stimulus pictures in a set configuration on a table in front of the participant. This flexibility in stimulus-presentation modality is important in allowing the EBI approach to be adapted to the particular needs of a given setting or participant. At the same time, table-top procedures carry the extra burden of requiring appropriate controls to prevent experimenter or teacher cuing (which could lead to false conclusions about emergent performance), and to ensure procedural integrity (the absence of which can dramatically decrease the effectiveness of teaching and testing methods). Tactics such as having the teacher always sit or stand behind the participant, and incorporating interobserver reliability measures, are common with these arrangements, for example.

Observing Responses

Control by the sample stimulus is critical to the acquisition of conditional discriminations in a matching-to-sample procedure. For this reason, on each matching trial, standard practice is to require the participant to make an *observing response* to the sample stimulus (e.g., a mouse click over the sample stimulus or a touch to the sample) once it is presented. The sample-observing response then results in presentation of the comparison stimuli and the opportunity to access reinforcers. The observing-response requirement is designed to increase the likelihood that the participant will actually notice (i.e., observe) the sample. In cases where sample control proves difficult to establish, requiring a unique response for each sample stimulus has often been effective (e.g., see Saunders & Williams, 1998). For example, a participant might be required to give a different but specific name for samples A1, A2, and A3 (e.g., Saunders & Spradlin, 1993; Pilgrim, Jackson, & Galizio, 2000); alternatively, a different gesture (e.g., a wave, a thumbs up, and a clap) or a different schedule performance (e.g., a fixed ratio [FR] and a differential reinforcement of low rates [DRL] schedule in a two-comparison procedure; Sidman et al., 1982) could be required in the presence of each sample. The advantage of a *differential sample requirement* is that it requires *successive discrimination* between the sample stimuli that are presented on different trials.

Similarly, requiring an observing response to each comparison stimulus prior to a final comparison selection can sometimes prove beneficial, especially if concerns arise about a participant making choices impulsively, without viewing each of the alternatives. For example, touching or naming each comparison in the array prior to the selection response should increase contact with each of the choices, and thereby increase the likelihood of control by the correct comparison.

Instructions

Task instructions in conditional discrimination training have varied from simple (e.g., "Touch one"; Lynch & Green, 1991) to more detailed (e.g., "Pick the one that matches" or "When you see this [A1], choose this [B1]"; Pilgrim, Jackson, & Galizio, 2000; Saunders, Saunders, Kirby, & Spradlin, 1988), and the best instruction set will again necessarily depend on instructional goals and the participants involved. For basic laboratory research in which the experimental question focuses on the role of the reinforcement contingency in generating equivalence, adding instructional control to the mix will complicate the analysis. Conversely, when producing an applied outcome is the sole focus and the isolation of controlling variables is a less important issue, a combination of instructions or other prompts and reinforcement contingencies can enhance acquisition of the conditional relations necessary to give rise to emergent performance. With some populations, such as young children or individuals with intellectual disabilities, reinforcement contingencies alone, in the absence of other programming, can be insufficient for establishing arbitrary conditional discriminations (e.g., Pilgrim, Jackson, & Galizio, 2000; Augustson & Dougher, 1992). As

long as the teacher ensures that the trained conditional discriminations are maintained once instructional prompts are removed, the groundwork is laid for equivalence outcomes.

Arranging Trial Types and Sequences

Careful attention to the balanced composition of trial types, as well as the order in which those trial types are presented within a block of trials, is essential to preventing a number of forms of competing stimulus control from interfering with conditional discrimination acquisition. Several standard practices have been developed specifically to guard against such issues. Let's say that we want to teach an AB conditional discrimination, such that comparison stimuli B1, B2, and B3 are selected in the presence of sample stimuli A1, A2, and A3, respectively. Three different trial types would be needed, one with each of the sample stimuli; the same three comparisons would be presented as choices with each sample. These three trial types should be presented the same number of times within a training trial block, to ensure an equal opportunity for reinforcement with each comparison. Imbalance here could create a bias toward selecting one (or more) comparison(s) over the others. Further, the position of the comparison stimuli should be counterbalanced across trials, such that each stimulus appears in each possible position on an equal number of trials. Imbalance here could result in a conditional relation between the sample stimulus and a particular position, rather than a particular comparison. Similarly, the position of the correct or S+ comparison stimulus should be counterbalanced across trials, such that the reinforced choice appears in each position on an equal number of trials. Failure to do so could again result in a preference for one stimulus location over the others. Finally, we would typically set a limit on the number of consecutive trials allowed within a block (e.g., no more than two or three trials in a row) (1) with the same sample stimulus, so that control by the sample is ensured, and (2) with the S+ comparison in the same position, so that responding to a particular position will not be inadvertently strengthened. With these constraints met, the order of trials within a block should otherwise vary unsystematically. To illustrate, the top half of Table 19.2 presents nine trials balanced with respect to the number of each trial type and the position of each comparison stimulus, including the reinforced or S+ choice. The bottom panel of Table 19.2 shows the same trials in a semi-randomized order that could prevent inadvertent reinforcement of sequences of stimulus or position selections. Consecutive trial blocks and/or teaching sessions should also use a newly randomized sequence of trials (maintaining the same constraints) so that a simple chain of responses cannot be established.

In sum, planning these details of trial arrangements prior to their presentation will help to ensure stimulus-control topography coherence—in this case, conditional control by the sample stimuli over specific comparison selections. Optimally, the number of trials presented in a trial block should be sufficient to allow for balance along each of the dimensions outlined above, while the number of trials or trial blocks to be completed within a single session will necessarily depend on the participant population. For example, young children may lose interest in the task after a relatively limited number of trials, so blocks ranging from 9 to 27 trials are common, while college students will often

TABLE 19.2 Arranging Balanced Trial-block Composition—Training

Sample Stimulus	Comparison Stimuli		
	Left Position	Center Position	Right Position
A1	***B1***	B2	B3
A1	B3	***B1***	B2
A1	B2	B3	***B1***
A2	B1	***B2***	B3
A2	B3	B1	***B2***
A2	***B2***	B3	B1
A3	B1	B2	***B3***
A3	***B3***	B1	B2
A3	B2	***B3***	B1
A3	***B3***	B1	B2
A2	B1	***B2***	B3
A1	***B1***	B2	B3
A3	B1	B2	***B3***
A2	***B2***	B3	B1
A3	B2	***B3***	B1
A1	B2	B3	***B1***
A1	B3	***B1***	B2
A2	B3	B1	***B2***

Note: ***Italicized boldface*** indicates the S+ or reinforced comparison selection.

work for periods of 1 to 3 hours and completion of many trial blocks is possible.

Instructional Programming

As noted above, for some participants or participant populations, teaching techniques in addition to the matching-to-sample reinforcement contingencies are necessary to produce mastery of the baseline arbitrary conditional discriminations. Verbal instructions have already been mentioned; other enhancements have included various prompting procedures, modeling the correct choices (e.g., Michael & Bernstein, 1991), and correction procedures, whereby a given trial is repeated until a correct choice is made (e.g., see Green & Saunders, 1998), for example. Some approaches have incorporated a systematic introduction of specific training-trial types designed to target particular dimensions of conditional control, in order to facilitate acquisition of the required baseline relations; these include teaching the component simple discriminations, stimulus-control shaping, exclusion training, and programmed introduction of the multiple conditional discriminations.

Teaching the Component Simple Discriminations. This training approach is based on the recognition that accurate conditional discrimination performances require two different types of simple discriminations, in addition to conditional control by the sample (e.g., Carter & Eckerman, 1975). These include *successive discrimination* between the sample stimuli, which are presented sequentially across trials, and *simultaneous discrimination* between the comparison stimuli, which are presented together as the choice array on a given trial. While the former often proves more challenging than the latter (e.g., Carter & Eckerman, 1975), assuming either can be problematic, particularly with young children or individuals with intellectual disabilities, and the absence of either will necessarily result in failure on the conditional discrimination task.

One teaching sequence designed to ensure all of the necessary component discriminations introduces a series of pretraining steps. Simple simultaneous discrimination is taught first with the stimuli that will serve as comparisons (e.g., B1, B2, and B3). Each stimulus serves as the S+ (that is, selection of that stimulus is reinforced) across phases (e.g., selecting B1, then B2, and then B3 is reinforced), and a new phase begins only after reliable selection of the S+ has been demonstrated (Saunders & Spradlin, 1993). In the second pretraining step, the stimuli that will eventually serve as samples (e.g., A1, A2, and A3) are presented individually, one at a time, over a series of trials, and a different response (e.g., naming, patterns of button pressing) is shaped and reinforced in the presence of each until this discriminative performance is strong (e.g., Saunders & Spradlin, 1989, 1993). In the third step, conditional discrimination trials are presented for the first time in blocks, such that sample A1 is presented for some number of trials in a row with B1, B2, and B3 as comparisons. A block of trials with sample A2, and then with A3, follows, with the same comparisons and for the same number of trials. The differential sample responses are still required, and the number of trials in a block is gradually decreased until the trial types with A1, A2, and A3 as sample are thoroughly intermixed. In this way, control by the sample stimulus alone becomes increasingly necessary. At this point, the procedure is a standard arbitrary matching-to-sample arrangement with accurate conditional responding established step by step.

Stimulus-control Shaping. Two approaches in this category resemble fading or errorless discrimination procedures by first teaching an easily acquired discrimination and then shifting gradually to a more difficult arbitrary conditional discrimination. In both cases, identity matching-to-sample serves as the starting point. Zygmont, Lazar, Dube, and McIlvane (1992) demonstrated the effectiveness of *stimulus-control shaping* with two 4- to 6-year-old typically developing children and two individuals with moderate to severe intellectual disabilities, none of whom had been successful in arbitrary match-to-sample training. The shaping programs began by first establishing identity matching with two abstract black-and-white-line drawings as sample and comparison stimuli; trials with each sample alternated here and throughout all training steps. Next, in Phase 1, the shape of one sample stimulus was modified in small increments across a series of training steps, while maintaining selection of the appropriate comparison. Figure 19.14 denotes the initial training stimuli as Step B for baseline, and shows that the sample and S+ (that is, the correct comparison) were physically identical at baseline; the subsequent rows of the figure illustrate the sequence of changes that were made to the sample drawing across training steps. When the target shape was achieved for Sample 1 (in Step 9, labeled FP for Final Performance), Phase 2 presented a similar training series for Sample 2, with gradual alterations made to its shape across steps. By the end of training, all four participants demonstrated accurate arbitrary conditional discrimination performance involving physically distinct samples and comparisons. Although the stimuli involved in this translational study were abstract line drawings, the shaping strategy could be easily applied to a wide range of teaching materials.

A second approach also shaped arbitrary sample control across a series of phases. To start, accurate identity matching (Pilgrim, Jackson, & Galizio, 2000) was established with young, typically developing children, first with pictures of familiar everyday objects (e.g., a pencil, a flower, and a boy's face), and then with abstract line drawings, indicating conditional control by sample stimuli across a range of stimulus types. The next training phase presented a *thematic matching* task in which the sample and correct comparison stimulus were physically dissimilar, but members of a common category or theme presumed to be familiar to the children. For example, with a picture of an ice cream cone as sample, selecting a picture of a birthday cake instead of a car or a bird was reinforced. Accurate performance on this training step extended conditional control to stimuli that had no physical resemblance to the correct comparison. When an arbitrary matching task with unfamiliar line drawings was presented next, the conditional discrimination was learned quickly, both by children who had failed to master the task previously and by experimentally naive children taught with this sequence from the start (Pilgrim et al., 2000).

Exclusion Training. This approach to teaching a new arbitrary conditional discrimination is based on a robust outcome known as **exclusion**, in which a novel comparison stimulus will be

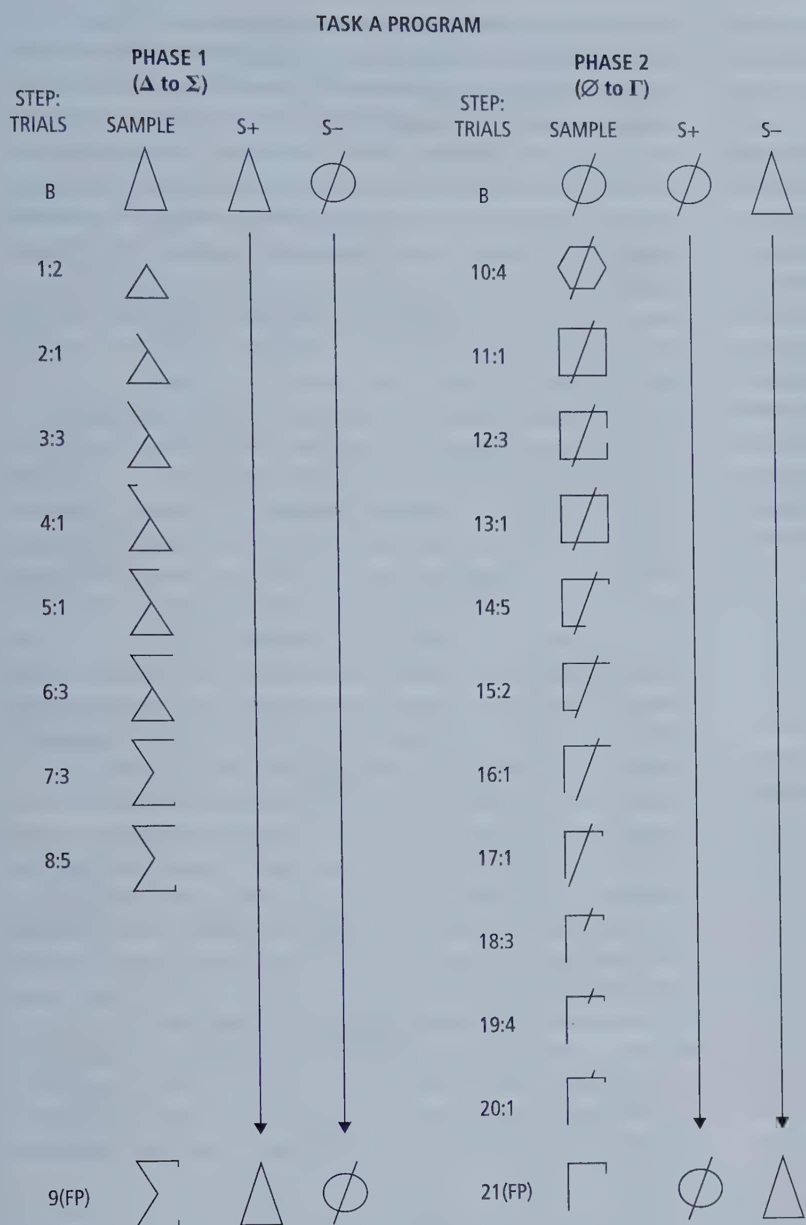


Figure 19.14 Steps of stimulus control shaping. Identity matching at baseline (B) was gradually changed to arbitrary matching by modifying the form of the sample stimulus in a series of steps, first for one sample (shown on the left side of the figure) and then for the other (shown on the right side of the figure).

selected over a known one in the presence of a novel sample (e.g., de Rose, de Sousa, & Hanna, 1996; McIlvane & Stoddard, 1981). Let's say, for example, that the starting repertoire of a young child with autism includes selecting the appropriate picture of a bird, a dog, and a cat when she hears the corresponding spoken word, but that matching printed letters of the alphabet to their spoken names occurs at chance levels of accuracy. A teacher might then present the spoken words "bird," "dog," and "C" as samples, with pictures of a bird, a dog, and the printed letter C as comparison choices. The exclusion effect would be demonstrated by the child's selecting the printed letter when presented with "C" as the sample, which would be followed by a reinforcer. Selecting other letters (e.g., B and D) in the presence of their spoken names could be established in similar fashion. Thus, by carefully intermixing new sample-comparison combinations with familiar stimulus relations, teachers can repeatedly create the opportunity for reinforcing target comparison selections in the presence of target sample stimuli. Although the name implies that the "known" comparison stimulus is excluded (a form of Type R or Reject relation), studies have consistently demonstrated that once demonstrated through exclusion, conditional control by the sample is maintained when the new relation is presented together with other samples and comparisons from the training set (rather than the already familiar combinations); that is, Type S or Select control is demonstrated. For example, a later training step following from our example above might present the three letter names as samples with the three printed letters as comparison choices. In this way, exclusion programming can be utilized to build the new arbitrary conditional discriminations required as baseline performances for equivalence outcomes.

Sequential Versus Simultaneous Introduction of Multiple Conditional Discriminations. The previous sections have focused on teaching a single conditional discrimination but, as described earlier, two (or more) conditional discriminations

(e.g., AB and AC) are necessary to allow for all defining tests of equivalence-class formation. Any of the teaching strategies outlined above could be applied in training additional conditional discriminations, of course, but fortunately, this is often unnecessary; establishing a first arbitrary relation frequently allows for rapid acquisition of those taught later (e.g., Pilgrim et al., 2000), an example of a learning-set outcome (Harlow, 1949). When children or individuals with intellectual challenges receive EBI, standard practice is to teach each conditional discrimination independently (e.g., AB training followed by AC training), and then to intermix the trial types making up each set of relations within a single trial block. A next conditional discrimination (e.g., AD) could then be trained in isolation, subsequently added to a mix of trials from all three, and so on, such that the baseline from which equivalence classes could emerge is developed sequentially. The alternative—introducing the mix of trial types from multiple conditional discriminations from the start of training—can sometimes save training time for select participant groups, such as college students.

Training Structures

Whether approached sequentially or simultaneously, teaching multiple conditional discriminations necessarily requires decisions about the specific sample-comparison combinations that will make up each conditional discrimination, and about the manner in which the different conditional discriminations will be interrelated. These dimensions of the teaching arrangement are described generally as the **training structure**. At least one stimulus set, often referred to as the **nodal stimulus** set or, more simply, the **node**, must be held in common across a minimum of two conditional discriminations to provide a basis for all equivalence properties. A number of different training structures are well represented in the experimental and applied literature on equivalence. For example, in Sidman's early studies reviewed in detail above, AB and AC conditional discriminations were trained. This arrangement is known as *one-to-many training*, in that one set of sample stimuli (e.g., the A stimuli in this case) is trained with multiple sets of comparison stimuli (e.g., B and C, in this example). A synonymous term for this training structure is *sample-as-node training*, which emphasizes that the A stimulus set serves a sample function in both of the trained conditional discriminations and thus provides the basis for emergent relations between stimuli from the different conditional discriminations (e.g., between B and C stimuli, in this case). A second popular training protocol is the *many-to-one* or *comparison-as-node* structure. In this arrangement, multiple sets of sample stimuli are trained with a single set of comparisons, and the comparison stimuli thus serve as node. For example, AB and CB conditional discriminations could be trained such that the B comparison stimulus set is held in common. A third structure is known as a *linear series training*. Here, the comparison stimuli from one conditional discrimination serve as sample stimuli in the next; for example, AB and BC relations might be taught. Each of these training structures could be applied with more than two conditional discriminations as well (e.g., AB, AC, and AD for one-to-many training; AB, CB, and DB for many-to-one training; AB, BC, and CD for linear series training, where two different nodes would be involved). Figure 19.15 illustrates these common training structures with a single stimulus from each of four sets (A to D). A total of three conditional discriminations are represented for each training structure, and the arrows in this diagram always point from a sample stimulus to the comparison selection that would be reinforced on a given training trial. In each case, teaching the matching performances shown by the arrows would serve as the baseline for emergent matching between any of the stimulus pairs not linked by an arrow. Finally, many examples of EBI have utilized mixed training structures with multiple nodal stimuli (e.g., AB, AC, and DC training); the variations in this category are extensive.

A number of studies have compared the effectiveness of these training structures for teaching the baseline conditional discriminations and for generating equivalence outcomes (e.g., Arntzen, 2012; Saunders, Saunders, Williams, & Spradlin, 1993). There are some reports of faster baseline acquisition with one-to-many training, perhaps because the more challenging successive discriminations between sample stimuli are required for only one set of stimuli (see Saunders & Green, 1999, for analysis), but these findings have not always been replicated. Similarly, across studies, one-to-many and many-to-one training structures appear to be equally effective for producing equivalence performance, with some studies favoring one-to-many slightly (e.g., Arntzen & Holth, 1997) and others favoring many-to-one (e.g., Fields et al., 1999). Certainly both provide a strong basis from which equivalence classes can be demonstrated. The one training structure that has consistently underperformed, compared to the others, in generating equivalence is linear series training (e.g., Arntzen, 2012), particularly as the number of trained conditional discriminations increases. Although linear series training has proved effective with many participants, this training arrangement is probably best avoided unless needed for other experimental reasons (e.g., to avoid a ceiling effect for equivalence successes, so that the impact of other procedural variables can be better analyzed). Either of the other structures

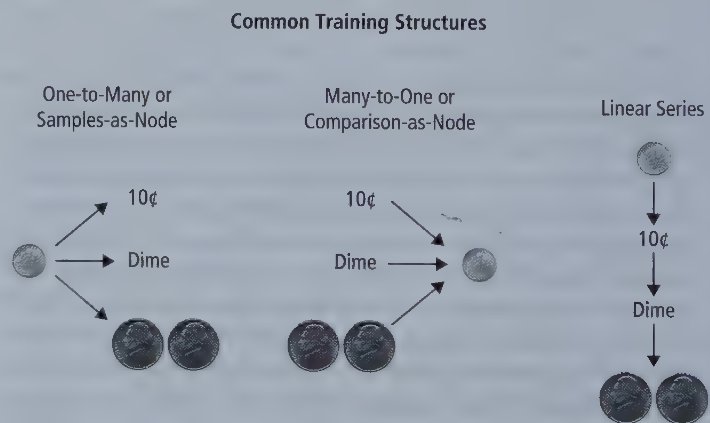


Figure 19.15 A schematic representation of three different training structures. Three conditional discriminations are represented for each structure, by presenting one stimulus from each of four sets (A, B, C, and D). The one-to-many (or sample-as-node) structure is presented in the left column, the many-to-one (or comparison-as-node) structure in the middle, and the linear series structure on the right.

can be a good choice for EBI, and the best selection will depend on factors such as the type of stimuli to be used (e.g., auditory stimuli are difficult to present as comparisons in the absence of special procedures) and the instructional goals of the EBI program.

Mastery (Accuracy) Criteria

The trained conditional discriminations provide the necessary prerequisites for equivalence-class formation, so it is essential that performance on all trial types is strong and consistent. To ensure this end, EBI programs incorporate predetermined *accuracy or mastery criteria* that must be met in each phase of training before moving on to the next. Typically, overall accuracy is measured for each trial block and defined as the number of trials on which the S+ comparison is chosen divided by the total number of trials in the block. (When acquisition is problematic, separate accuracy measures for specific trial types—e.g., all trials with each sample stimulus—are important to analyze as well.) Because intermediate accuracy scores (e.g., 75%) can reflect multiple sources of stimulus control (e.g., see Sidman, 1980, 1987), the mastery criteria are typically set at a relatively high accuracy percentage and are often required for more than one consecutive trial block (unless the number of trials composing the block is large). For example, in one series of studies (Pilgrim & Galizio, 1990; Pilgrim, Chambers, & Galizio, 1995), adults and children received conditional discrimination training in trial blocks consisting of 16 trials each. The mastery criteria for each training phase required accuracy of at least 88% (14 out of 16 trials correct per block), with no more than one error on any given trial type, for two consecutive sessions.

Scheduling Consequences During Training

Match-to-sample training for one or more conditional discriminations will typically include reinforcers for selection of the S+ comparison on every trial until the specified mastery criteria are met. Choosing an appropriate and effective reinforcer for EBI training is subject to the same considerations that would hold for any applied behavior analysis program. Common examples have included points or tokens worth money or prizes, small edibles, written feedback (e.g., "Correct" or "Incorrect"), attractive visual displays or cartoons, and auditory jingles or chimes. Some teaching programs also include mild punishers (e.g., a brief buzzer, a short time-out, or a trial repetition) for S− selections, again programmed for every training trial until mastery is achieved. Before moving immediately from continuous reinforcement to tests for the properties of equivalence, however, a common practice is to first arrange intermittent reinforcement, where the number of trials per block that include programmed consequences is decreased, often in a series of steps, with mastery criteria required at each step. For example, the steps might include having reinforcers available on 100%, then 75%, and then 50% of the trials. At issue in these reinforcement reductions is preparation for the introduction of the test or *probe trials* designed to evaluate equivalence-class formation. Probe trials are typically presented in extinction, to allow for interpretations of emergent versus directly trained relations. Without adequate preparation, the absence of reinforcers for novel probe performances can produce variability in comparison selections, whereas training to mastery with intermittent reinforcement helps to ensure response maintenance in the absence of trial-by-trial consequences. Relatedly, one benefit of including punishers as part of the training contingency is that doing so provides a basis for discrimination between incorrect responses and the mere absence of a reinforcer.

Some EBI programs have included final training steps with still leaner schedules, arranging programmed consequences for only 20% or even 0% of the trials (e.g., Sidman & Tailby, 1982). Although motivation can become an issue, experience with lean reinforcement schedules for trained relations makes discrimination between training trials and later probe trials less likely, with the net result that probe responding is not differentially extinguished. Strategies designed to counteract the low reinforcer density have included providing alternative bases for reinforcement (e.g., color-matching trials presented at the end of the AB and AC training block; e.g., Sidman & Tailby, 1982) and/or instructions (e.g., "In this part of the game, I can't tell you whether you are right or not, but you should continue to do your best."). As usual, the appropriateness of such strategies will depend on the participant population and the reasons for conducting the EBI program.

Testing Protocols

As was the case for baseline training, a number of different testing procedures are reflected in the equivalence literature. Many decisions about testing arrangements carry important advantages and disadvantages, such that careful consideration is warranted for procedural details, including the presentation and order of probe trials, consequence scheduling, and criteria for equivalence-class formation.

Composition of Probe-trial Blocks

Two general methods are used for presenting probe trials: *massed testing* and *interspersing* probe trials with baseline-trial types. In the first method, the trial block is composed of probe trials only. In the second, the block includes probe trials as well as each of the baseline training trial types presented in a semi-randomized order, often with no more than one or two probe trials presented in succession. Massed testing allows for more rapid assessment of equivalence-class formation, but when equivalence outcomes are not demonstrated, this approach leaves open the question of whether the prerequisite baseline relations have been maintained. Introducing novel trial types in discrimination studies can sometimes disrupt previously mastered performances, at least temporarily, and this issue may be particularly relevant with some participant populations. If the trained conditional discriminations are not strong, there is no reason to expect equivalence to emerge, and in the absence of baseline-trial types, this possibility can be neither evaluated nor remedied. Related strategies have included beginning test sessions with a baseline-review trial block (e.g., Kennedy, Itkonen, & Lindquist, 1994) or alternating baseline-review and massed-probe trial blocks. These arrangements necessarily mean that the equivalence assessment is less rapid and actually represents approximations of the interspersed trial-type approach. In any case, some provision for evaluating the

maintenance of baseline performances is critical in those cases in which equivalence outcomes are not demonstrated. When interspersed, the proportion of probe trials to baseline trials varies considerably across EBI programs and is often determined by the overall reinforcement density being targeted (i.e., with no reinforcers for any probe trial, and intermittent reinforcement for the baseline trials) and the necessary balanced composition of trial types described for training arrangements. To illustrate, Table 19.3 presents a balanced trial block composed of three symmetry probes and nine baseline training trials.

Scheduling Consequences During Testing

When baseline-trial types have been included as part of the test sessions, decisions must be made about scheduling consequences for those trials. In some EBI programs, consequences are programmed for at least a portion of the baseline trials (e.g., Sidman et al., 1985), such that the trained relations are not only evaluated but also possibly strengthened during testing, and overall task motivation is facilitated. In other work, extinction conditions are in effect for all trial types—baselines as well as probes (e.g., Bush et al., 1989). The latter arrangement is often preceded by instructions, as described previously for leaning the reinforcement density during training (e.g., "You will no longer receive feedback on your selections, but there is still a correct answer, and you should continue to do your best;" e.g., Dougher et al., 1994), although the effectiveness of this strategy will be limited with participants who have poor verbal skills. An advantage of testing in complete extinction is that at least one basis for discrimination between baseline- and probe-trial types is removed, and the possibility that probe trials will be treated differentially (e.g., attended to less carefully) is thereby lessened.

There is also the issue of programming consequences for probe trials. Certainly, the overwhelming majority of equivalence-based instruction, basic or applied, has presented all probe trials under extinction conditions, because a primary research question involves the extent to which new or emergent performances result from the experimental training. In an alternative strategy, however, reinforcers are arranged for probe responses consistent with equivalence-class formation (e.g., Kastak, Schusterman, & Kastak, 2001). In these cases, data analysis is focused on comparison selections for the first presentation of each probe-trial type. Because direct reinforcement cannot yet be responsible for these responses, conclusions about emergent equivalence are still possible. Further, when EBI is implemented for purely practical reasons—for example, to establish an important functional skill for the first time—answering the research question noted above may be less important than ensuring optimal conditions for the acquisition of that skill (e.g., Sidman, 1981), and these will surely include reinforcement of the skill when it is demonstrated, for whatever reason.

Test Order

Whether probe trials are massed or interspersed with baseline trials, one common practice in EBI is to present only one probe type (i.e., reflexivity, symmetry, transitivity and/or combined equivalence) per test-trial block. This practice can allow for multiple presentations of each probe combination within a test block, and it simplifies data analysis. As with the other procedural variables discussed in this section, however, the literature reveals great variety in the order of presentation of the probe types. Indeed, the range of test approaches has run the gamut from introducing the most involved test types first (e.g., after training AB, BC, and CD conditional discriminations, presenting DA tests for combined equivalence as the first probe types; Sidman & Tailby, 1982; Sidman et al., 1985) to beginning with the "simplest" (e.g., testing for BA symmetry first, after AB training combination; Imam, 2006). As we've also seen for the

TABLE 19.3 Arranging Balanced Trial-block Composition—Interspersed Probe and Baseline Trials

Sample Stimulus	Comparison Stimuli		
	Left Position	Center Position	Right Position
A3	*B3*	B1	B2
A2	B1	*B2*	B3
B1*	**A2**	**A3**	**A1**
A1	*B1*	B2	B3
A3*	B1	B2	*B3*
A2	*B2*	B3	B1
B2*	**A1**	***A2***	**A3**
A3	B2	*B3*	B1
A1*	B2	B3	*B1*
B3*	***A3***	**A1**	**A2**
A1	B3	*B1*	B2
A2*	B3	B1	*B2*

Notes: Italics indicates the S+ comparison stimuli. **Boldface** indicates probe trials.

* indicates a trial with no programmed consequences.

other procedural variables, the optimal sequence of test presentation will depend largely on the goals of the particular study and the participant population.

For programs of EBI with young children or individuals with intellectual disabilities, where the generation of equivalence classes is of primary interest (as opposed to the factors underlying them), a **simple-to-complex** testing protocol has much to recommend it (e.g., Adams, Fields, & Verhave, 1993). In this arrangement, probe types are introduced sequentially, beginning with symmetry, followed by transitivity (if relevant), and then combined tests for equivalence. Similarly, probe types requiring fewer nodes are presented prior to those requiring more nodes.[4] Further, each probe type is introduced immediately following the minimal training necessary to provide the prerequisites for the probe relation; when probes and baseline trials are interspersed, only those baselines that serve as prerequisites for the probed relation would be included.

For example, consider an EBI program that teaches three conditional discriminations—AB, BC, and CD, as in the linear series training illustrated in Figure 19.15. The simple-to-complex testing protocol for this training is outlined in Table 19.4, and involves the following sequence of training and testing phases: train the AB (coin to price sign) conditional discrimination to mastery; present BA (price sign to coin) symmetry probes interspersed with AB baseline trials; train BC (price sign to written coin name) to mastery; present CB (written coin name to price sign) symmetry tests interspersed with BC baseline trials; train mixed AB and BC; present AC (coin to written coin name) transitivity tests, and then CA (written coin name to coin) combined tests in that order, both interspersed with AB and BC baseline trials; train CD (written coin name to quantity of nickels) to mastery; present DC (quantity of nickels to written coin name) symmetry tests interspersed with CD baseline trials; train mixed BC and CD; present one-node BD (price sign to nickels) transitivity tests and then one-node DB (nickels to price sign) combined tests, both interspersed with BC and CD baseline trials; train mixed AB, BC, and CD; present two-node AD (coin to nickels) transitivity tests and then two-node DA (nickel to coin) combined tests, both interspersed with AB, BC, and CD baseline trials. This approach to systematic increases in test complexity develops experience with each of the necessary simple discriminations as it progresses, and has been shown to increase the likelihood of strong equivalence outcomes (e.g., Adams et al., 1993). For older typically developing children and adults, fewer steps will be necessary, but a testing sequence of symmetry followed by transitivity and combined tests at each nodal level can maximize class formation.

Judging Class Formation

Our earlier section titled "Class Formation" described the basics for assessing this most fundamental EBI outcome. At issue is whether emergent performances indicative of stimulus interchangeability will be demonstrated. Because interpretations of equivalence require reliable class-consistent responding on the defining probe-trial types (e.g., Sidman & Tailby, 1982), data analysis typically focuses on the percentage of probe trials on which comparison selections reflect the equivalence classes that would be expected to follow from the training provided. For example, if teaching signed word to picture (AB) and signed word to written word (AC) conditional discriminations produced equivalence relations, A1B1C1 and A2B2C2 classes would be predicted, with each class including the corresponding signed word, picture, and printed word. For trials presenting a printed-word C1 as sample, then, class-consistent responses would include selections of the matching signed word, picture, or written word (A1, B1, or C1) comparisons on symmetry, combined, and reflexivity tests, respectively. The percentage of class-consistent responses could thus be calculated, either for

TABLE 19.4 A Simple-to-Complex Testing Protocol

Training/Testing Phase	Training Trial Types	Test-Trial Types
1	AB	
2	AB	BA
3	BC	
4	BC	CB
5	AB & AC	
6	AB & AC	AC
7	AB & AC	CA
8	CD	
9	CD	DC
10	BC & CD	
11	BC & CD	BD
12	BC & CD	DB
13	AB, BC, & CD	
14	AB, BC, & CD	AD
15	AB, BC, & CD	DA

all probe trials combined or, especially when class-consistent selections are below 100%, for each probe type individually. When multiple test blocks are presented (versus one massed test presentation, for example), these measures are often graphed and followed across the individual tests. As was the case with judging baseline acquisition, intermediate percentages of class-consistent responses (e.g., 60% to 80%) are indicative of multiple sources of stimulus control (e.g., see Saunders & Williams, 1998; Sidman, 1980, 1987). For this reason, strong conclusions about class formation require relatively high levels of class-consistent responding (e.g., 90% or greater). As with many of our procedural variables, the specific criteria vary across studies, and for purely practical uses, any increase in the probed relations may provide a useful basis for further development.

When the expected emergent performances are not demonstrated immediately, several different strategies have proven helpful. Perhaps the simplest involves repeated presentation of the probe blocks. As noted previously, delayed emergence, in which the probe performances improve with additional testing, is a relatively common finding (e.g., Sidman et al., 1985; Sidman et al., 1986). Delayed emergence has been interpreted as a weakening in competing sources of stimulus control that are not available consistently across probe trials, leaving the equivalence relation as the dominant source of control over comparison selections (e.g., Sidman, 1992, 1994). In addition to repeated testing, low scores on initial probe tests are often improved by re-exposure to the training blocks, especially those that include the baseline prerequisites for the problematic probe types. Other intervention strategies include training or testing modifications that target particular forms of competing stimulus control. For example, if a probe-trial error analysis documents a position preference in which a participant consistently chooses the lower right corner of the computer screen regardless of the stimulus presented, the presentation format could be redesigned so that no stimulus appears in that position on any trial.

Procedural Variations

Teaching a minimum of two conditional discriminations in the match-to-sample format provides the core prerequisites for assessing equivalence-class formation. However, a number of alternative training approaches have been utilized effectively in efforts to (1) explore the range of procedures that might generate equivalence classes; (2) enhance the efficiency of equivalence-based instruction; and (3) enhance the variety of performances that might be produced by EBI. To illustrate, a few of the most prominent of these alternatives are reviewed here, including the use of compound stimuli, class-specific reinforcement, and simple discrimination training.

Compound Stimuli

One procedural variation builds on the matching-to-sample training format by presenting a stimulus compound as the sample. For example, the printed word DOG and a picture of a dog could be presented one above the other as a pair in the sample-stimulus position, and selecting a particular comparison stimulus (e.g., the printed word CANINE) would be reinforced in the presence of that compound. EBI programs utilizing such compounds have revealed that each component of the compound can come to function as an independent equivalence-class member; that is, in the present example, DOG, CANINE, and the dog picture could all be matched interchangeably when presented individually (e.g., Groskreutz, Karsina, Miguel, & Groskreutz, 2010; Lane & Critchfield, 1998). Thus, by using a compound sample, a given amount of training results in many more emergent relations and larger equivalence classes—a finding with important implications for EBI.

In one demonstration of the applied significance of this finding, Stromer and MacKay (1992) used compound stimuli in match-to-sample training designed to target spelling performances in three 9- to 13-year-old boys with academic deficits. Each sample stimulus included a printed word together with a picture (e.g., DOG and a line drawing of a dog). In addition to using compound sample stimuli, the procedure incorporated three additional variations from the standard teaching approach described thus far. First, an *identity match-to-sample* procedure was employed; that is, the S+ comparison stimulus was physically identical to one part of the sample compound. In the present case, on any given trial, the comparison stimuli included either three different pictures, one of which was identical to the picture component of the compound sample, or a group of 10 letters, some combination of which would be identical to the written word component of the sample. The two types of identity trials (i.e., with picture or letter comparisons) were presented equally often in an unpredictable order. Selection of the comparisons that were physically identical to the sample resulted in reinforcement. The second procedural variation involved in this study is known as a *constructed response requirement*. On trials that presented letters as comparisons, the participant's task was to touch, in the correct order, the letters that produced the written word from the sample compound. After the identity matching performances had been taught, a third variation was introduced—a *delayed match-to-sample procedure*. In this step, the compound sample was presented as usual to start each trial; however, an observing response to the compound sample removed it from the computer screen, and the comparison stimuli were presented without the sample being present. Because the comparisons to be presented were unpredictable (i.e., pictures versus letters), this addition required attention to both components of the sample.

The top half of Figure 19.16 illustrates the delay procedure; each column represents a different trial type as it would have appeared on the computer screen. In each case, the compound sample is shown in the top box, and an observing response to the sample produced the display illustrated in the bottom box. Note that the sample stimulus is no longer present in the bottom display, and that the comparisons to choose from include either three pictures for the trial type shown on the left or, for the trial type shown on the right, an array of letters from which the sample word could be composed.

Over the course of the study, this training was arranged with three different sets of printed words (i.e., dog, cat, owl; canine, feline, avian; etc.), one set at a time. The same pictures appeared as part of the sample compound with all printed-word sets; that is, the pictures were held in common across the trained conditional discriminations and thus could serve as nodes. The bottom half of Figure 19.16 illustrates training trials with words from the second and third word set, on the top and bottom rows,

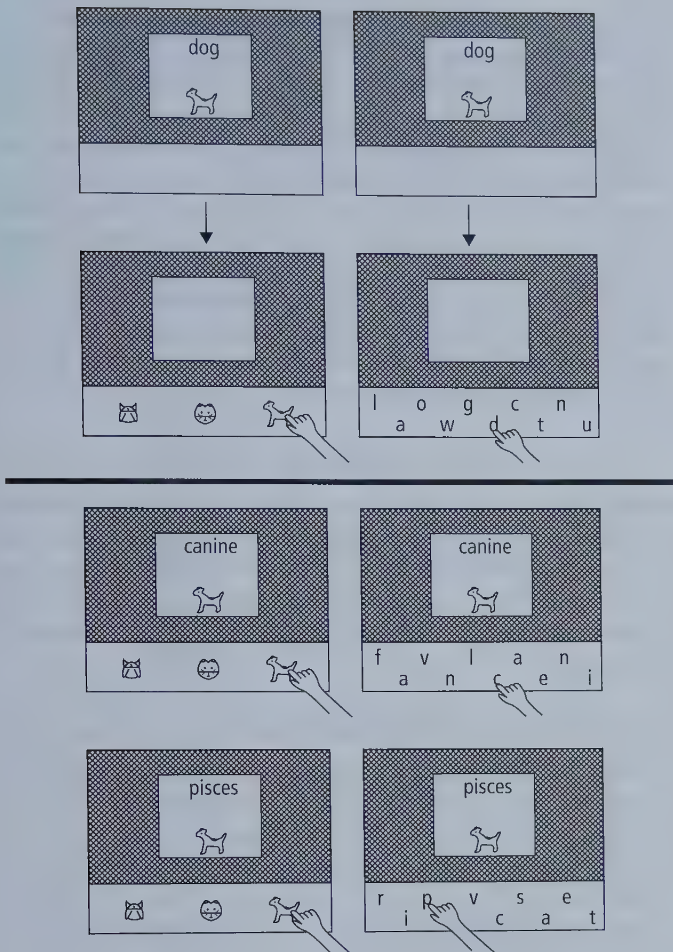


Figure 19.16 The top half of the figure illustrates the delayed matching task. Each column represents a different trial, with the compound sample shown first; an observing response to the sample produced the display in the bottom box. The bottom half of the figure illustrates training trials with words from the second and third word sets. The delay procedure is not represented here but was in place.

Adapted from "Spelling and Emergent Picture-printed Word Relations Established with Delayed Identity Matching to Complex Samples," by R. Stromer and H. A. Mackay, 1992, *Journal of Applied Behavior Analysis*, *25*, p. 898.

respectively; the delay procedure was in place here, too, but is not represented in this figure. Note that in these trial examples, the picture of the dog is common to each sample compound.

Test trials presented each word and picture individually as samples; all possible combinations of the printed words, pictures, and constructed responses were tested after training with each of the three stimulus sets (i.e., a multiple baseline across stimulus sets design). For each test, all three boys demonstrated emergent performances involving only those stimuli for which training had been provided. By the end of training, probe performances revealed the formation of equivalence classes, each of which included a picture, three different printed words, and three different constructed words. That is, any of the visual stimuli within a class could be matched to any other (e.g., DOG and CANINE could be matched) and could occasion any of the three word constructions (e.g., with CANINE as sample, the word DOG could be constructed, depending on the letter choices provided). Further, when the corresponding spoken words were presented as samples on probe trials, any of the three printed words or the picture in the class was selected, and the appropriate spelling was constructed. At this point, the boys had an established spelling repertoire, resulting from EBI.

Class-specific Reinforcement

A second variation on the match-to-sample format involves a particular arrangement of consequences known as *differential outcomes* or **class-specific reinforcement**. In this procedure, not only is the correct comparison choice conditional on the sample stimulus, but the type of consequence delivered is, too. For example, given sample A1, correct selection of comparison B1 would produce Reinforcer 1: a small marshmallow, perhaps. Given sample A2, correct selection of B2 would produce a different reinforcer: Reinforcer 2—let's say, a sip of juice. Teaching a second conditional discrimination would work similarly; for example, selecting comparison C1 given sample A1 would produce the marshmallow, while selecting comparison C2 given sample A2 would produce access to juice. Finally, let's add identity match-to-sample training with D stimuli, such that selecting D1 given a D1 sample produces the marshmallow, with selecting D2 given a D2 sample produces juice.

Several important findings have been obtained with this procedure. Not only can differential outcomes facilitate the acquisition of conditional discriminations (e.g., Litt & Schreibman, 1981), but *reinforcer-probe trials* reveal that the class-specific consequences themselves become members of the equivalence classes, thereby increasing the number of class members that can be added by teaching a single conditional discrimination (e.g., Dube, McIlvane, Mackay, & Stoddard, 1987; Dube, McIlvane, Maguire, Mackay, & Stoddard, 1989). From our example above, when pictures of the marshmallow or juice are presented as sample or comparison stimuli, they are matched with the related A, B, C, or D stimuli (i.e., these are the reinforcer probes). Further, the A, B, C, and D stimuli prove to be interchangeable with each other, indicating equivalence-class formation (e.g., Dube et al., 1987, 1989). Thus, the class-specific reinforcer can also serve as a node for class formation; in this example, the D stimuli could have become class members only by virtue of their relation to the class-specific consequences held in common across the trained relations.

These sorts of findings also carry important implications for equivalence-based instruction. First, differential-outcomes procedures provide another route to producing more emergent relations with a given amount of training, because the reinforcers join the respective classes. Second, equivalence-class formation can result from identity match-to-sample training when class-specific reinforcers are utilized (e.g., Dube et al., 1987, 1989), and for certain target populations, identity training is more easily mastered than the arbitrary training that is more typical in equivalence research (e.g., Pilgrim, Jackson, & Galizio, 1990; Zygmont et al., 1992).

Studies are beginning to explore class-specific reinforcement procedures as a route to increasing the power of equivalence-based applications. To illustrate, preschool children who could count but otherwise scored poorly on math pretests were first taught to match

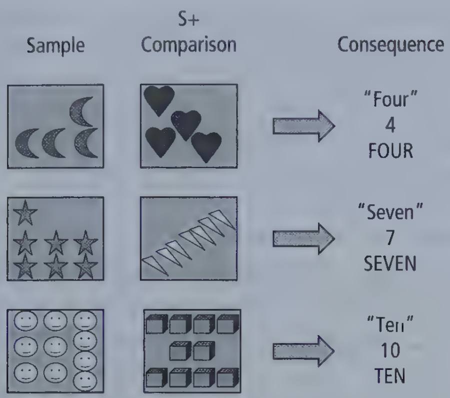


Figure 19.17 An illustration of the trained conditional discrimination. Each row represents a trained relation, with sample stimuli presented in the left column, the S+ comparison in the middle column, and the compound consequences that followed selection of the S+ in the right column. Note that the consequence for any given trial included the spoken word and either the numeral or the written word.

pictures with equal quantities of objects totaling 4, 7, or 10 (e.g., given 4 moons as sample, selecting 4 hearts rather than 7 lightning bolts or 10 squares; Luffman, 2012). Incorrect matches produced a brief buzzer, while correct matches produced a two-part class-specific consequence. Every correct match produced the corresponding spoken quantity value (e.g., "Four!") and one of two visual stimuli, the corresponding numeral (e.g., 4) or the written word (e.g., FOUR), each presented on half of the relevant trials. Figure 19.17 summarizes the training; each row of the figure presents an example of the matches that were taught, with the sample stimulus presented in the left column, the correct comparison in the middle column, and the possible consequences in the right column.

On the probe trials that followed this teaching, the children demonstrated class-consistent matching of each quantity, numeral, written word, and spoken word. Later teaching phases arranged the same compound consequences for accurately matching quantities of objects to a "+ 1" addition problem (e.g., given four objects, selecting "3 + 1," rather than "6 + 1" or "9 + 1"; AB training) and then for matching quantities of objects to a "+ 2" addition problem (e.g., given four objects, selecting "2 + 2," rather than "5 + 2" or "8 + 2"; AC training). Figure 19.18 shows all of the stimuli presented in AB and AC conditional discrimination training, including the class-specific consequences.

After completing the three teaching phases, probe performances documented the emergence of three equivalence classes with at least six members (i.e., a quantity of objects, the corresponding numeral, spoken and written words,

Stimulus Sets Used in Conditional Discrimination Training

A Quantities Trained	B + 1 Stimuli	C + 2 Stimuli	Consequences
4	3 + 1	2 + 2	"Four," 4, FOUR
7	6 + 1	5 + 2	"Seven," 7, SEVEN
10	9 + 1	8 + 2	"Ten," 10, TEN

Figure 19.18 Stimulus sets used in conditional discrimination training. Note that the consequence for any given trial included the spoken word and either the numeral or the written word.

"+ 1" and "+ 2" addition problems), any of which could be matched to any other interchangeably by children who had been unable to do so prior to training. (Probe batteries were presented just before and just after each training phase.) In this example, EBI resulted in the symbolic relations defining of a rudimentary but fundamental start to understanding mathematics.

Three-term Contingency Training

As noted previously, the acquisition of arbitrary conditional discriminations has sometimes proved challenging for young children and individuals with developmental delays. Thus, for both practical and theoretical reasons (Sidman, 1994, 2000), the possibility of generating equivalence classes based on **simple discrimination**, or three-term contingency, training has been explored, and with promising results (e.g., Debert, Matos, & McIlvane, 2007; Pilgrim, Boye, Hogan, & Groff, in press; Sidman, Wynne, Maquire, & Barnes, 1989). Indeed, the ease with which simple discriminations are typically acquired suggests considerable potential for this approach to applications of EBI.

To illustrate, consider a follow-up to the math teaching program described above. In this case, class-specific consequences were arranged again, but for simple discrimination training (Yonkers, 2012). Four young children with learning disabilities were identified by their teachers as needing extra help in basic math skills. The first training phase taught three simple discriminations with the correctly oriented numerals 4, 7, or 10 as the S+ stimuli. For each discrimination, the two S− stimuli presented the S+ numeral in either an upside-down or a 90° orientation. Figure 19.19 illustrates a trial presentation for each of the three discriminations, with the S+ and two S− stimuli appearing in three of the four corners of the computer screen.

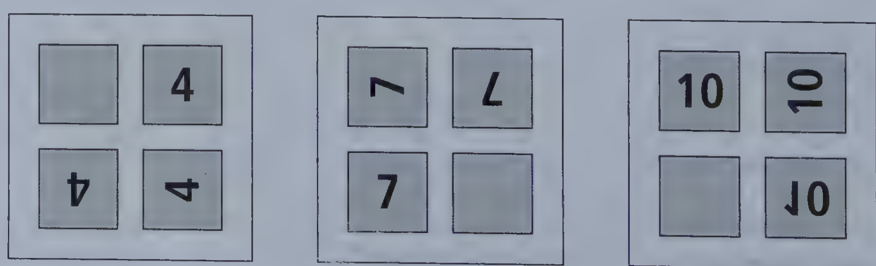


Figure 19.19 An illustration of example training-trial configurations, one with each of three trained simple discriminations.

Selecting either S− stimulus produced a buzzer, while selecting the S+ produced a compound class-specific consequence consisting of the spoken numeral name (e.g., "Four!") and the corresponding written word (e.g., FOUR), as illustrated in Figure 19.20.

Rapid acquisition of these simple discriminations was followed by conditional discrimination probe trials in which the spoken word, the written word, and the numerals were matched class-consistently, regardless of which served as sample and

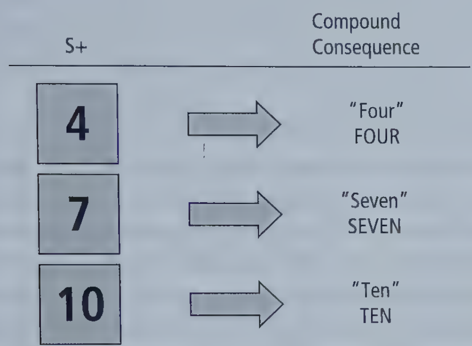


Figure 19.20 An illustration of the trained simple discriminations. Each row represents the S+ stimulus in the left column, and the compound consequences that followed selection of the S+ in the right column.

which as comparison. Additional simple discriminations were taught in subsequent training phases, in which the stimuli presented depended on each child's skill level, as measured by pretests. For the children with the lowest pretest scores, the B stimulus set consisted of correctly oriented "+ 1" addition problems (i.e., 3 + 1, 6 + 1, and 9 + 1), while the C stimulus set consisted of correctly oriented "+ 2" problems (i.e., 2 + 2, 5 + 2, and 7 + 2). One training trial of each type is illustrated in Figure 19.21; Figure 19.22 presents the S+ stimuli for each of the nine trained simple discriminations.

Again, selecting one of the correctly oriented S+ stimuli, rather than an upside-down or sideways presentation of the problem, resulted in the same class-specific consequences used for numeral training (i.e., the appropriate spoken and written number name, also illustrated in Figures 19.21 and 19.22). For children with more advanced skills, multiplication problems or Roman numerals were presented as stimulus sets, illustrating the flexibility of this approach to meet individual children's needs. Upon completion of the simple discrimination training with differential outcomes, match-to-sample probe tests documented the formation of equivalence classes (e.g., 4, "Four", FOUR, 3 + 1, and 2 + 2). Figure 19.23 illustrates just a few of the many probe arrangements made possible by the training program, with a sample stimulus presented in the center of the screen and comparison stimuli presented in three of the four corners of the screen. In this

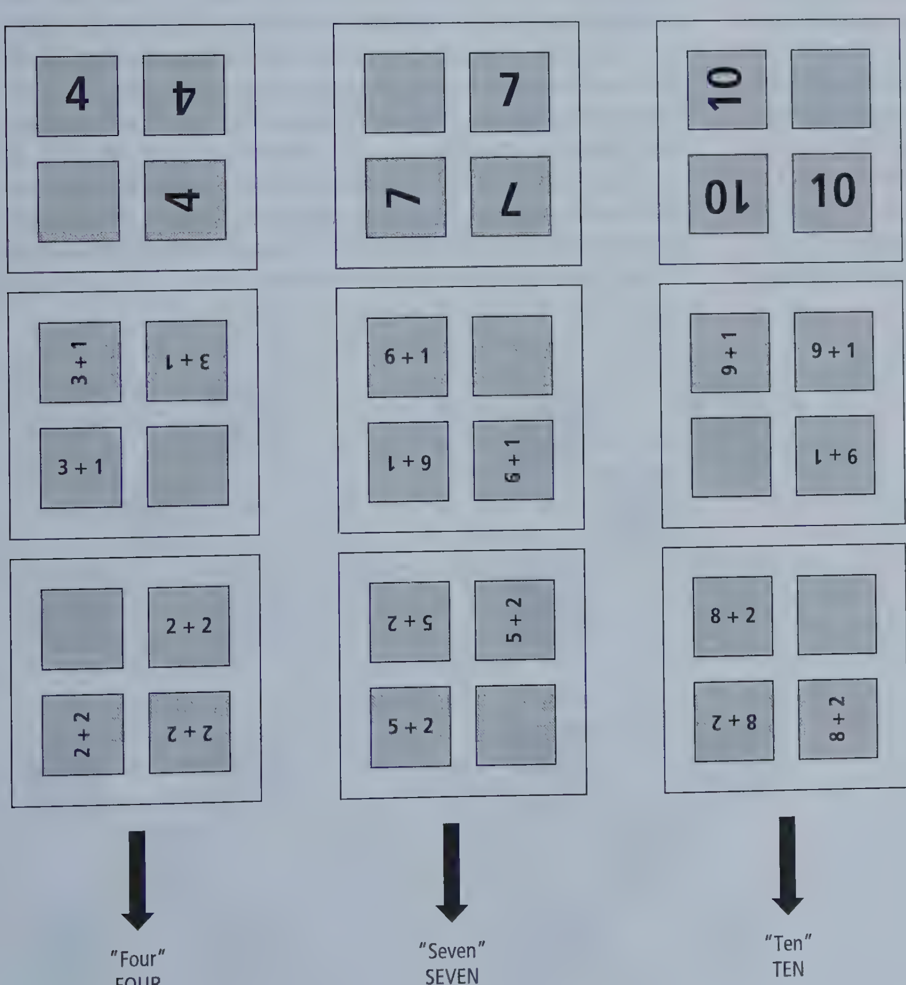


Figure 19.21 Illustration of a trial presentation with each of the nine trained simple discriminations. Each row represents trials with a particular stimulus set (A, B, or C), and each column represents trial types for which the same class-specific compound consequence was programmed, as indicated in the bottom rows.

Stimulus Sets Used in Simple Discrimination Training

A	B	C	Consequences
Numerals	+ 1 Stimuli	+ 2 Stimuli	
4	3 + 1	2 + 2	"Four," FOUR
7	6 + 1	5 + 2	"Seven," SEVEN
10	9 + 1	8 + 2	"Ten," TEN

Figure 19.22 A representation of all of the S+ and consequence stimuli presented during simple discrimination training.

EBI program, then, nine rapidly learned simple discriminations resulted in three five-member equivalence classes, as evidenced by more than 60 emergent conditional performances. These findings not only established the feasibility of such an approach to EBI but also provided a striking demonstration of the efficiency of EBI for producing generative and functional symbolic skills.

APPLICATIONS AND GENERALITY

The past years have witnessed enthusiastic attention to the possibility that equivalence-based instruction might be utilized creatively to establish or enhance performances across varied participant populations and across an increasingly broad range of target skills. The inherent flexibility of the equivalence paradigm, in which literally any stimulus combination can be employed, and the increasing variety of effective teaching and testing procedures that it has inspired, make the approach particularly attractive to applied behavior analysts. As each new application is documented, the generality of the equivalence outcome is further confirmed, a critical issue in the development of any scientific concept. In the sections that follow, some highlights from the accumulating application literature will be reviewed, to illustrate the scope and the promise of the equivalence approach to instruction with children and adults with intellectual disabilities, young typically developing children, adult clinical populations, and college students.

Children and Adults with Intellectual Disabilities and Young Typically Developing Children

To date, these populations stand as the major recipients of equivalence-based applications, because skill acquisition is a necessary and ongoing focus for such individuals. Important categories of EBI focus have included language- and number-skills training as well as other functional academic and life skills.

Language-Skills Training

Perhaps the greatest advances in programmatic EBI are found in the area of language-skills training. Building on Sidman's (1971; Sidman & Cresson, 1973) original findings, a number of studies have replicated the formation of equivalence classes as a basis for establishing rudimentary reading skills (e.g., Elias, Goyos, Saunders, & Saunders, 2008, with adults with intellectual disabilities; Osborne & Gatch, 1989, with hearing-impaired preschool children; Sidman, Cresson, & Willson-Morris, 1974, with two boys with Down syndrome), where each class included a spoken word or sign (demonstrated in American Sign Language), a written word, and a picture or object. Participants, all of whom were nonreaders at the start, went on to sign or orally name both the pictures/objects and the written words following class formation.

As illustrated earlier in this chapter, another group of studies has demonstrated that spelling performances can be effectively incorporated into an equivalence approach (e.g., de Souza et al., 2009; Goyos, Souza, Silvares, & Saunders, 2007; Melchiori, de Souza, & de Rose, 2000). The constructed-response procedure has often been used for this purpose whereby, in the presence of a spoken or written word, the participant's task is to produce that word by selecting a series of letters or syllables from a choice pool, in

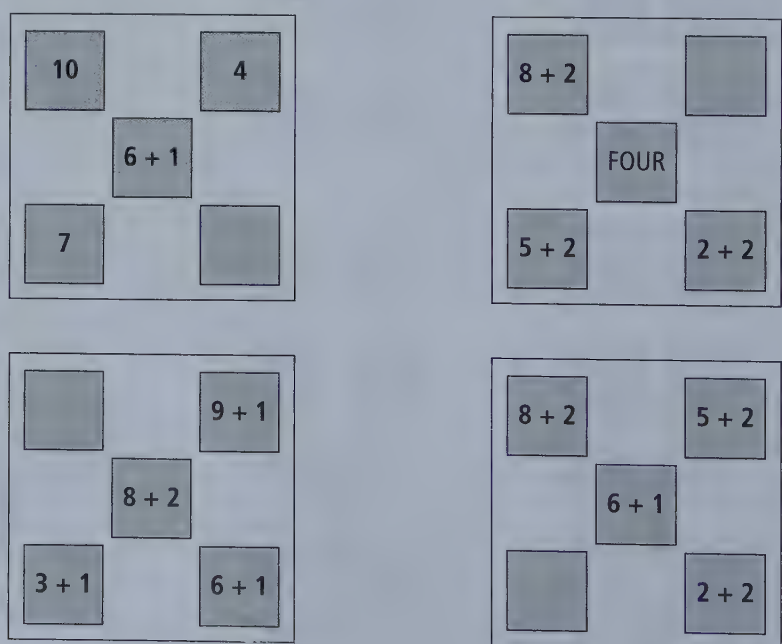


Figure 19.23 Examples of conditional discrimination test-trial arrangements.

order, either by touching each choice sequentially on a computer presentation or by moving letter tiles in a table-top task. This teaching approach has proven effective with a range of participant groups including preschoolers, young students from regular and from special education classes, and adult nonreaders.

One particularly exciting outcome of the studies just cited has been that spelling improvements successfully generalized to words that were not directly targeted in the teaching steps. As was the case in the work described above with compound sample stimuli, it seems that the individual components, or units, of the trained compounds came to exert stimulus control on their own (e.g., the individual letters, in the case of spelling). Thus, when these smaller units appeared in novel arrangements (e.g., when new words were presented), appropriate responses (e.g., accurate spelling sequences) were possible. The ability to produce and recombine minimal units (e.g., words, syllables, phonemes, letters) in the presence of novel stimulus compounds is clearly essential to many functional language skills, such as reading and spelling. Not surprisingly, then, another important development in equivalence-based language-instruction programs has involved incorporating steps to deliberately facilitate such recombinations, also known as *recombinative generalization*. The general strategy involves teaching with compounds that have been specifically chosen to include overlap in the individual units that will be necessary for correct responding to novel compounds. For example, a child might be taught to select the written words MAP, SAP, MUM, or SET correctly upon hearing the corresponding spoken word, and then prove able to select the novel words MET or SUM on test trials. This teaching arrangement is sometimes described as *matrix training*, to emphasize the designed overlap in units across training and testing compounds. The key is to establish all of the prerequisite units for successful performance prior to test trials with the new combinations. (For further elaboration and examples, the interested reader should consult, as representative sources, de Souza et al., 2009; Hanna et al., 2011; Mueller, Olmi, & Saunders, 2000).

One group of behavior analysts has extended the generality of this training strategy considerably with a sophisticated, comprehensive, and ongoing reading-instruction program for children in Brazil who have experienced repeated failure in the public school system (de Souza et al., 2009). Many of the children are from families with low socioeconomic status and they carry significant risk for school drop-out. The reading curriculum as a whole incorporates many of the findings described in this chapter as contributing to maximal effectiveness, and it has produced many first-time readers as a result (e.g., de Rose, de Souza, Rossito, & de Rose, 1992; Matos & Hubner-D'Oliveira, 1992). By way of overview, some key characteristics of the early instructional phases include a matching-to-sample procedure with constructed responses and a matrix training approach that builds on an interesting feature of the Portuguese language. Many Portuguese words are formed by consonant-vowel syllables (e.g., BOLA, which means "ball" in English, and LATA, which means "can"). Recombinations of these syllables form new words (e.g., BOTA, which means "boot" in English), thus setting the stage nicely for training designed to produce generative reading repertoires, with syllables as the minimal units. For the constructed responses, students learn to select the appropriate sequence of syllables (rather than individual letters, as in the spelling studies above) given a printed word, dictated word, and picture. AB (spoken word to picture matching) and AC (spoken word to written word matching) conditional discriminations are also taught, as in several of our earlier examples (e.g., Sidman, 1971).

Tests at this point reveal the formation of equivalence classes (i.e., spoken word, picture, and written word) both for trained words and for recombinative generalization words. Significant improvements have also been obtained in reading and in spelling, both in the syllable-constructed response format and in cursive writing, again with training and generalization words alike. Subsequent teaching phases of this integrated instructional program move from acquisition of reading and writing syllables to whole words to sentences and short stories, and refinements continue. This increasingly complete reading curriculum has now benefited more than 3000 students at reading centers in multiple schools throughout Brazil, and may well stand as the single most comprehensive application of equivalence-based instruction in existence to date.

Number-Skills Instruction

Just as basic reading comprehension has been approached through EBI, so has basic numeric comprehension, with targets ranging from simple number classes to more complex mathematical concepts. Examples of the former have involved match-to-sample training in which participants learn to select the appropriate quantity of objects or pictures (e.g., *, **, or ***) in the presence of dictated number names (e.g., "one," "two," and "three"), and the appropriate numeral (e.g., 1, 2, or 3) in the presence of the same number names (i.e., AB and AC conditional discriminations). Probe tests following such training have reliably resulted in evidence of equivalence-class formation, with each class including a dictated name, a quantity, and a numeral (e.g., Gast, VanBiervliet, & Spradlin, 1979; Maydak, Stromer, Mackay, & Stoddard, 1995; both with individuals with intellectual disabilities). Emergent matching of quantities and numerals supports an interpretation of number comprehension, and participants come to name the quantities and the numerals correctly as well. Variations on this training have also established number names from two unfamiliar languages (i.e., Native American Ojibwe and Dakota words) as equivalence-class members (Haegele, McComas, Dixon, & Burns, 2011) with typically developing preschoolers.

These simple number classes have been elaborated in several ways in subsequent studies. One example was described previously (Luffman, 2012) in which match-to-sample training using compound class-specific reinforcers was conducted with young typically developing children, and probe performances documented the formation of six-member equivalence classes that included quantities, numerals, spoken number names, written number names, and $X + 1$ and $X + 2$ arithmetic problems for each of three values (i.e., 4, 7, and 10; see Figure 19.18 for all stimulus sets). Each class member was matched accurately to every other class member, yielding hundreds of emergent numeric relations, after teaching only AA, AB, and AC conditional discriminations. The children in this study also learned to perform a constructed response. For each quantity sample presented during training, they were required to slide the

appropriate number of beads on an abacus, prior to the presentation of comparison stimuli; that is, a class-specific constructed-observing response was required. At the study's conclusion, generalization trials were presented in which the three numerals served as either samples or comparisons, and novel addition problems equaling 4, 7, or 10 (e.g., 2 + 2, 4 + 3, 5 + 5) served as comparisons or samples, respectively. (Note that, as illustrated in Figure 19.18, all values from 1 through 10 had been presented during training). The children made accurate matches on these novel trial types, and they performed the appropriate constructed response to the novel sample combinations, thus demonstrating recombinative generalization of the values.

Many beginning math learners struggle to understand and respond accurately with respect to fractions and decimals, and EBI has been designed to facilitate these more complex mathematical concepts as well (e.g., Hall, DeBernardis, & Reiss, 2006; Lynch & Cuvo, 1995). To illustrate, match-to-sample training was conducted with seven 11- to 13-year-old students identified as having math difficulties; grade-equivalent math scores for these fifth- and sixth-graders ranged from 3.0 to 3.7 (Lynch & Cuvo, 1995). AB and then BC conditional discriminations were taught in which the A stimuli included 12 fractions expressed as ratios (e.g., 3/5), the B stimuli included 12 fractions represented pictorially (i.e., as a grid of 100 squares with the appropriate number of squares shaded), and the C stimuli included 12 values presented as decimals (e.g., 0.60). The top portion of Figure 19.24 illustrates the trained and tested relations. In contrast to generally poor pretest scores, all seven students improved dramatically, often to near-perfect levels, on the fraction-to-decimal probe tests (i.e., AC and CA) indicative of equivalence-class formation. Thus, 12 equivalence classes consisting of a fraction, a shaded grid, and a decimal value emerged, and a subsequent AD training step with 12 new fractions (D stimuli) expanded the classes to include a fourth interchangeable member. Results were mixed on paper-and-pencil tests for solving fractions with no choices provided, and on match-to-sample generalization tests involving novel fractions, with some of the students showing large improvements from pretests and others none at all. It might be noted that the teaching procedures in this early study did not utilize a matrix-training approach.

Another obviously important functional-skill category for young children and individuals with disabilities involves money equivalencies, in which dictated or printed prices and various coin or dollar combinations must become interchangeable (see our example from Figure 19.7). This target has also been the subject of a number of EBI studies (e.g., Keintz et al., 2011; Stoddard, Brown, Hurlbert, Manoli, & McIlvane, 1989) using the match-to-sample procedure. In an early classic report (McDonagh, McIlvane, & Stoddard, 1984), a woman with moderate intellectual disability was taught to match two different coin combinations to a printed price for each of three values (5, 10, and 15 cents). With no further training, she could then match the coin combinations to each other and give the values of each, demonstrating equivalence-class formation. The three-member classes were expanded successfully by teaching additional coin combinations for some of the values, and these results were replicated systematically across multiple individuals with intellectual disabilities (e.g., Stoddard et al., 1989). In a more recent illustration (Keintz et al., 2011), two preschoolers with autism received training with three dictated coin names (the A stimuli; "penny," "nickel," and "dime"), three actual coins (the B stimuli; a penny, a nickel, and a dime), three printed prices (the C stimuli; \$0.01, \$0.05, and \$0.10), and three dictated prices (the D stimuli; "1 cent," "5 cents," and "10 cents"). After learning AB, BC, and DC conditional discriminations (see the bottom half of Figure 19.24), both children demonstrated four-member equivalence classes including the dictated coin and price names, the actual coins, and the written prices, and both named each of the coins. Collectively, the results from these studies demonstrate the great potential of EBI for establishing symbolic equivalencies involving large numbers of individual coin relations.

Other Functional Skills

A range of skills in addition to those involving basic language and number classes have been established through EBI, providing further testament to the considerable flexibility of this teaching approach. For children and adults with cognitive disabilities and typically developing children, several examples will serve to reflect the variety of performances that have been targeted to meet specific needs of the individual participants.

One interesting study (Arntzen, Halstadtrø, Bjerke, & Halstadtrø, 2010) focused on facilitating music skills for a 16-year-old with autism and intellectual disability; the boy was interested in music and could play a few simple songs on a keyboard when guided by a color-code that showed which note to play. The equivalence-based training was designed to teach him to read and play musical chords. For each potential class, the stimulus sets in this study included (a) the name of a chord in Norwegian (the boy's second language); (b) a keyboard diagram with dots indicating finger placement; (c) a musical staff with the chord represented as notes; and (d) the name of the chord in Vietnamese (the boy's first language). Across the phases of the study, eight different major chords and an equal number of minor chords were targeted; none of these chords were played or named correctly on a pretest. Posttest results verified the emergence of a large number of four-member equivalence classes. The authors also reported that the boy continued to use the chords after the study concluded, and they recommended EBI approaches for establishing other hobby skills, to possibly enhance the lives of individuals with intellectual disabilities. A similar study with typically developing preschoolers was also effective in establishing simple music skills with an equivalence-based approach (Tommis & Fazey, 1999). Again, improvements in playing simple melodies were maintained, at a 7-week follow-up test in this case.

Several studies have focused on expanding sources of stimulus control for key functional skills in young children and adults with disabilities. For example, activity schedules are an important aid to helping children work independently in many special education classrooms (e.g., MacDuff, Krantz, & McClannahan, 1993; McClannahan, MacDuff, & Krantz, 2002). Typically, activity schedules present a sequence of pictures, one at a time, each of which indicates a particular task to be completed. To make the developmentally appropriate shift from control by pictures to control by written words, an equivalence-teaching approach was arranged for two 6-year-old children with autism who responded accurately to picture schedules (Miguel, Yang, Finn, & Ahern,

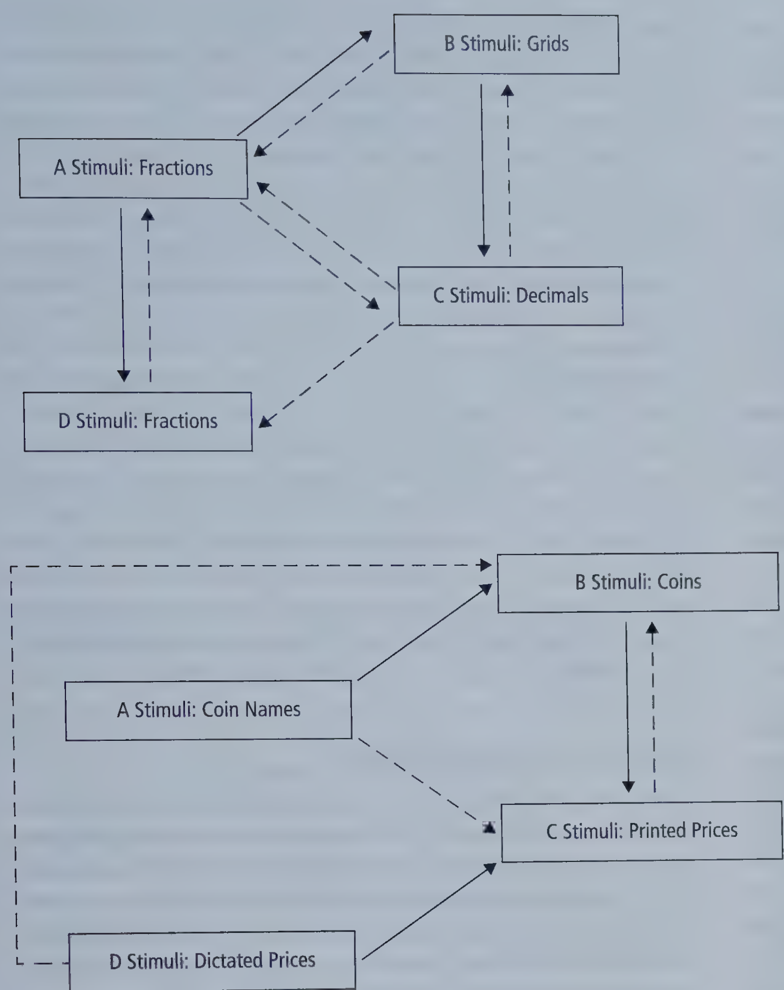


Figure 19.24 The top half of the figure presents a schematic representation of trained (solid arrows) and tested (dashed arrows) conditional discriminations from Lynch and Cuvo (1995). The bottom half of the figure presents the same information from Keintz, Miguel, Kao, and Finn (2011).

2009). After being taught to match six activity-schedule pictures to their dictated names (AB training) and six corresponding written words to the same dictated names (AC training), both children demonstrated accurate responding to activity schedules with written words, as well as accurate matching of words and pictures, indicating comprehension of the words, and oral reading. A similar strategy was used (Rehfeldt & Root, 2005; Rosales & Rehfeldt, 2007) to augment the manding skills of male adults with severe intellectual disabilities who had learned to request items with pictures via the Picture Exchange Communication System (PECS) (Frost & Bondy, 1994). After successful picture-exchange requests had been established, participants were taught to choose the correct picture given its dictated name and to choose the corresponding written word given the same dictated names (AB and AC conditional discriminations). Although unable to do so on pretests, all participants demonstrated equivalence-class formation, matching pictures to their written words and vice versa, and they successfully requested items with the appropriate written word. Some of the men also named the written words and requested items by name as well as by written-word exchanges. For the participants who received maintenance tests 1 month after training (Rosales & Rehfeldt, 2007), lasting improvements to functional communication were demonstrated.

Equivalence-based instruction has also been used to teach academic content such as geography to children with disabilities, as well as to typically developing children. For example, basic geography lessons were presented to two boys with autism by teaching them to match a drawing of each of nine states to the dictated state name (AB training) and then to match the name of the state's capital city to the state drawing (BC training; LeBlanc, Miguel, Cummings, Goldsmith, & Carr, 2003). Where all pretest scores were at approximately chance levels of accuracy for both boys, posttests revealed strong emergent performances indicative of equivalence-class formation; state names were matched to their capitals and vice versa, and each trained relation was reversible. One of the boys also demonstrated perfect accuracy on vocal posttests in which correct responses had to be produced orally (rather than selected) in response to relevant geography questions about the states involved in training (e.g., "What is the capital of Florida?"). This teaching approach was replicated systematically in a study with five adolescents with fragile X syndrome where three state names, drawings, and capitals were taught (Hall et al., 2006). In another variation, also involving geography content, five typically developing 6-year-olds were taught AB and BC intraverbal sequences rather than match-to-sample performances (Perez-Gonzalez, Herszlikowicz, & Williams, 2008). In this procedure, a geography question was presented orally (e.g., "What is the capital of Argentina?"), and the child was taught to give the correct answer orally. Thus, the procedure involved production of an oral response rather than the

more typical selection of a stimulus in a match-to-sample task (i.e., topography-based rather than selection-based responding). The stimulus sets included (a) two countries; (b) two capital cities; and (c) two city parks. After mastering the trained intraverbal sequences (i.e., country-capital and capital-park), probe trials presented novel questions with each possible combination of elements (i.e., capital-country, park-capital, country-park, and park-country). Results showed accurate emergent responding on at least some probe types for most participants.

Another important research direction for EBI in applied behavior analysis involves utilizing equivalence classes as an explicit means to promote generalization after teaching other functional skills. That is, if a new performance is taught in the presence of one stimulus setting, and that stimulus is a member of an equivalence class, we might expect the performance to be controlled by other members of the class as well (i.e., the transfer-of-function effect). Although this possibility has received relatively little attention in applied work, some initial exploration has been promising. Taylor and O'Reilly (2000) worked with six participants with mild intellectual disabilities to establish a response chain of grocery-store shopping skills. A task sequence (e.g., enter the store, consult a two-item shopping list, find the first item, place it in the cart, etc.) was trained first in a classroom, and then in a particular grocery store for all six. For two participants, this constituted the complete training (i.e., single-instance training). For another two participants, the same task-sequence training was conducted in two additional grocery stores (i.e., multiple-exemplar training). For the remaining two participants, single-instance training was followed by EBI. AB and BC conditional discriminations were taught with spoken words ("Supermarket," "Shop," and "Restaurant"), the corresponding written words, and photographs taken from inside each setting. An interesting feature of the training was that three different photographs of supermarket settings were presented across training trials. Probe trials for these participants documented the formation of three equivalence classes, each including a spoken word, a written word, and the relevant photographs. After training, shopping skills were evaluated for all participants in a generalization grocery setting. The task sequence was followed accurately and equally well for the participants who received EBI and multiple-exemplar training, while those who received only single-instance training performed more poorly. Although this study is limited by the small number of participants in each teaching condition, the results suggest that EBI can provide the same positive impact on generalization as more costly and labor-intensive training in multiple sites. Certainly, these are findings worthy of further exploration.

A final and significant category of equivalence-based instruction with children involves programs designed to facilitate skills related to sensory deficits, such as vision or hearing. One illustrative study used an equivalence approach to teach Braille letter naming to four children (ages 7 to 12 years) with degenerative visual impairments (Toussaint & Tiger, 2010). All four students could name printed letters, but despite having a year or more of classroom Braille instruction, none demonstrated proficiency in reading Braille words. The stimulus sets used for teaching and testing included (a) Braille letters; (b) printed letters; and (c) spoken letter names, for 5 to 26 letters, depending on the participant. With printed letter-name relations (BC relations) already intact, teaching focused on having the children match Braille letters to printed ones (AB relations). On subsequent posttests, all four children demonstrated strong emergent matching between printed-letter samples and Braille comparisons, and between spoken name and Braille comparisons. They also named the Braille letters, verifying the formation of equivalence classes that provided a basis for further Braille reading instruction. With respect to hearing deficits, another program of study has undertaken the goal of establishing symbolic function for auditory stimulation provided by a cochlear implant in children with deafness (Almeida-Verbu et al., 2008; da Silva, de Souza, & de Rose, 2006). To illustrate, 6- to 9-year-old children with early-onset deafness first received AB and AC conditional discrimination training with all visual stimuli (Almeida-Verbu, et al., 2008). After three-member equivalence classes were demonstrated, a new conditional discrimination was taught with auditory stimuli as samples (DC training). All participants then demonstrated class expansion, indicating that the sounds had become the symbolic equivalent of the existing class members (e.g., D1 was interchangeable with A1, B1, and C1, and so on). These preliminary results suggest great promise for helping young children learn the meaning of sounds once cochlear implants make auditory stimulation available to them, perhaps for the first time in their lives.

Adult Clinical Populations

Although a relatively small category compared to many we have reviewed, applications of EBI to adult clinical issues have been explored in several studies, noteworthy because of the potential that is exemplified for these populations. As a case in point, individuals with acquired brain injury or illness often display deficits for which equivalence approaches seem well suited, such as the inability to provide or respond to words for everyday objects, events, or people. One study (Cowley et al., 1992) focused on three adult men with brain injury who were unable to relate their therapist's name to his or her face, even after many months of working together—a problem for independent functioning within the rehabilitation facility. The solid arrows in Figure 19.25 depict the stimulus relations that were demonstrated by each of the men at pretest for each of three different therapists. The relations targeted in training are depicted by the bold arrow, and the potentially emergent relations are indicated by the broken arrows. Posttest results revealed equivalence-class formation for all three men; that is, spoken names, written names, and faces were matched together in any combination. The two men who received the test were also successful at providing oral names for their therapists' faces. Facial-emotion recognition was targeted in another study with three adults with brain injury (Guercio, Podolska-Schroeder, & Rehfeldt, 2003). Although these participants were unable to do so on pretests, learning to match two sets of facial expressions to their appropriate written labels (i.e., happiness, sadness, and anger; AB and CB training) produced significant improvements in matching different representations of the same emotion to each other, and in providing correct oral labels for those emotions.

The field of gerontology represents another obvious target for equivalence approaches, although work in this area is in its infancy. Although it is clear that some elderly individuals demonstrate equivalence-class formation readily (e.g., Perez-Gonzalez & Moreno-Sierra, 1999; Saunders, Chaney, & Marquis, 2005), it is also apparent that age matters (Wilson & Milan, 1995), with

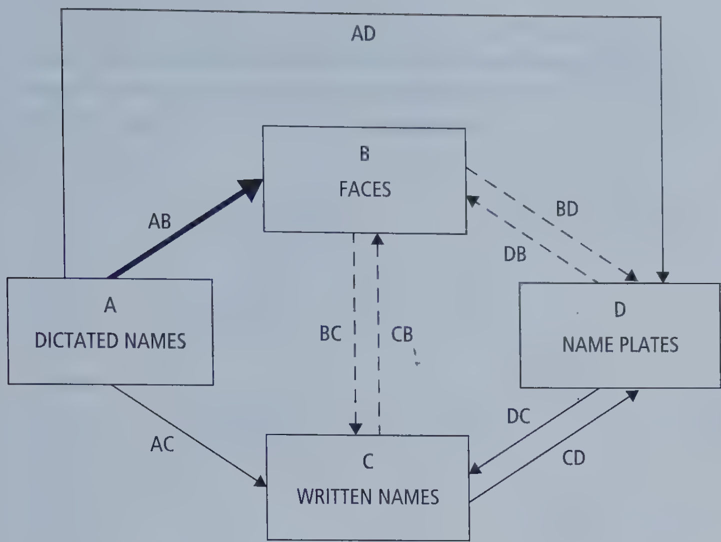


Figure 19.25 A schematic representation of the existing (thin solid arrows), trained (bold solid arrow), and tested (dashed arrows) conditional discriminations.

Using Stimulus Equivalence Procedures to Teach Name-face Matching to Adults with Brain Injuries" by B.J. Cowley, G. Green, and B. Brounling-McMorrow, 1992, Reproduced with permission of John Wiley & Sons Inc.

older participants (62 to 81 years old) less likely to demonstrate class formation than young adults (19 to 22 years old). Further, a correlation has been noted (Gallagher & Keenan, 2015) between demonstrations of equivalence-class formation and scores on the Mini Mental Status Examination (MMSE), a widely used assessment of cognitive impairment in the elderly. Indeed, the ability to demonstrate equivalence-class formation may prove to be a more sensitive measure of impairment, in that deficits in class formation were present for individuals with MMSE scores above those indicative of cognitive difficulties. One investigation of a participant diagnosed with Alzheimer's disease (an 84-year-old woman with an MMSE score of 20, or mild cognitive impairment) found evidence of delay-identity matching but not equivalence-class formation (e.g., Steingrimsdottir & Arntzen, 2011; see also Steingrimsdottir & Arntzen, 2014). Equivalence approaches may thus have much to offer in the way of assessing cognitive declines. Whether EBI can be used effectively to target lost skills, or to facilitate maintenance of remaining skills at least in individuals with mild impairment, remains to be explored.

The broad domain of health care would seem to hold many possibilities for skill improvement based on EBI. In one innovative study relevant to the critical issue of obesity (Hausman, Borrero, Fisher, & Kahng, 2014), researchers sought to teach college students to make accurate estimations of portion size because portion size can be directly related to caloric intake, especially for overweight individuals (e.g., Burger, Fisher, & Johnson, 2011). Nine students, five of whom were overweight or obese, were taught to match measuring cups filled at a given level (i.e., one-quarter cup, one-half cup, and one or two cups of a dry food) to a quantity of the same food piled on a plate, and vice versa (i.e., AB and BA training). They were also taught to match one of four measuring aids (i.e., a golf ball, tennis ball, baseball, and softball) to each measuring cup amount, and vice versa (i.e., BC and CB training). Figure 19.26 illustrates these training procedures. At pretest, posttest, and 1-week maintenance and generalization tests, a large plastic container filled with a food was presented, and each participant was asked to move one-half cup of the food to a plate. No cues, aids, or measuring cups were provided. The generalization tests involved novel foods that had not been presented during training sessions.

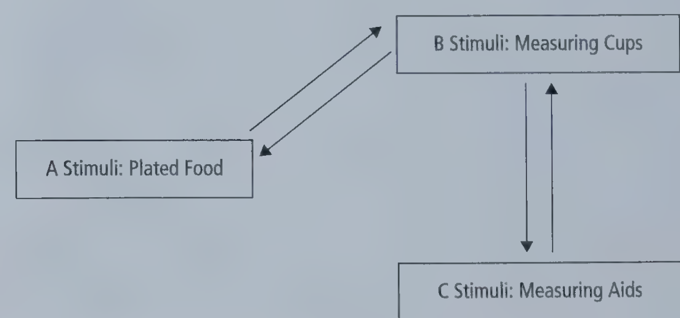


Figure 19.26 A schematic representation of the trained conditional discriminations.

All nine participants demonstrated improvements in the accuracy of their portion estimates from pretraining to posttraining and for seven, those improvements were maintained 1 week later. Five participants demonstrated similar accuracy with at least one novel food in the generalization tests. Thus, a brief teaching intervention (50 to 90 trials) was sufficient to impact a potentially influential skill relevant to healthful eating practices. Further investigation of EBI for health-related skills is called for.

In the College Classroom

The benefits to be gained from programmed instructional techniques based on behavioral principles have long been noted for university-level instruction (e.g., Keller, 1968, 1982; Skinner, 1968), and modern technology has only increased the feasibility of providing course content in a more effective manner. Of course, many college classes or course sequences require facility with large sets of basic facts and relations before advanced study is possible. Unfortunately, the rote memorization of facts that characterizes some common study techniques (e.g., flash cards) often fails to suffice for true understanding of the subject. Instead, comprehension of the sort at issue throughout this chapter is necessary for progress. It only stands to reason, then, that behavior analysts would apply equivalence-based instruction to teaching university-course content, extending the generality of the instructional approach to much more complex stimulus sets, a sophisticated group of learners, and new settings.

One early study of this type (Fields et al., 2009) sought to teach introductory psychology students concepts of statistical interaction in a two-way analysis of variance, a subject matter that has presented challenges for many learners at this level. Match-to-sample training established AB, BC, and CD conditional discriminations with the four sets of stimuli depicted in Figure 19.27, and probe performances documented the emergence of four classes, each with four

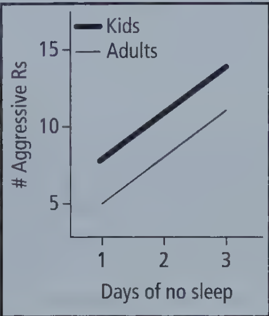
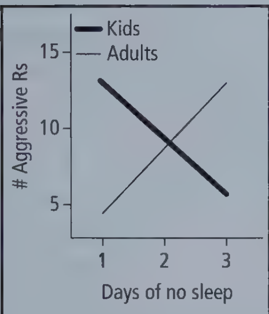
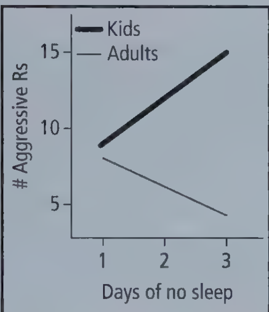
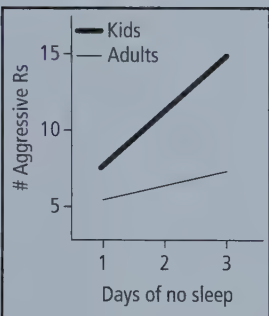
A1	B1	C1	D1
	Aggression was directly related to sleep deprivation for both kids and adults. For each level of sleep deprivation, a decrease in age produced a constant increase in aggression. Degree of aggression in kids and adults did not intersect at any level of sleep deprivation.	No Interaction	Independent variable A produces the same directional change in responding for all values of independent variable B. Changing the value of independent variable B produces a constant change in the effect of each level of independent variable A. The effects do not intersect.
A2	B2	C2	D2
	In kids, aggression was inversely related to sleep deprivation. In adults, aggression was directly related to sleep deprivation. Age reversed the effect of sleep deprivation on aggression. Levels of aggression seen with kids and adults intersected at an intermediate level of sleep deprivation.	Crossover Interaction	Changing the value of independent variable B reverses the effects of independent variable A, and these effects intersect.
A3	B3	C3	D3
	In kids, aggression was directly related to sleep deprivation. In adults, aggression was inversely related to sleep deprivation. Age reversed the effect of sleep deprivation on aggression. Levels of aggression in kids and adults did not intersect at any level of sleep deprivation.	Divergent Interaction	Changing the value of independent variable B reverses the effects of independent variable A, and these effects do not intersect.
A4	B4	C4	D4
	Aggression was directly related to sleep deprivation for both kids and adults. Increased sleep deprivation produced a larger increase in aggression in kids than in adults. Levels of aggression in kids and adults did not intersect at any level of sleep deprivation.	Synergistic Interaction	Independent variable A produces the same directional change in responding for all values of independent variable B. Changing the value of independent variable B enhances the effect of independent variable A. The effects do not intersect.

Figure 19.27 Stimulus sets used in training AB, BC, and CD conditional discriminations. Each column represents a stimulus set (A, B, C, D), and each row represents a potential equivalence class.

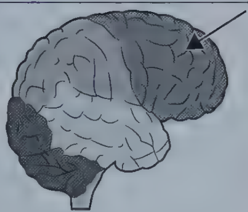
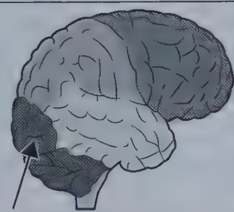
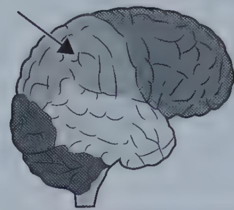
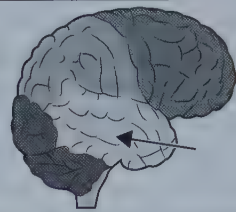
From "Equivalence Class Formation: A Method for Teaching Statistical Interactions" by L. Fields, R. Travis, D. Roy, E. Yadlovker, L. de Agular-Rocha, and P. Sturmey, 2009, Reproduced with permission of John Wiley & Sons Inc.

members. Each class included a line graph illustrating a particular statistical interaction, text describing the interaction shown in the graph, the label for the type of interaction, and text defining the interaction type. Interchangeability between the members of each of the four classes (i.e., no interaction, crossover interaction, divergent interaction, and synergistic interaction) indicated comprehension of the statistical concepts, based solely on the EBI. As further evidence, the students completed a paper-and-pencil multiple-choice test of the same interaction types but with novel graphs and graph descriptions, both before and after the match-to-sample training. After completing the 90-minute training, students showed significant improvement in test scores, from an average of approximately 50% on the pretest to an average greater than 90% on the posttest; all students earned scores corresponding to a grade of "A" or "B." Thus, the comprehension that emerged from training generalized to new exemplars. In contrast, a matched group of control students took the test twice with 90 minutes in between, but did not receive training. For these students, there was no change in test scores as a result of repetition, and 73% of them earned scores corresponding to a grade of "D" or "F."

Fienup, Covey, and Critchfield (2010) arranged an EBI program designed to take advantage of the generative power of class expansion or merger. The lesson content in this study was brain structures and functions, another complex subject matter consisting of a great number of interrelated facts. Volunteer undergraduate students received match-to-sample training with the stimuli illustrated in the top half of Figure 19.28. All possible transitive/equivalence relations were pretested and posttested following each phase of training. The diagram at the bottom of Figure 19.28 represents all tested relations with dashed lines and all trained relations with solid lines. The first phase of training established AB and AC conditional discriminations, and probe tests documented the formation of four three-member classes (i.e., ABC classes, each including the name of a lobe of the brain and two different functions corresponding to that lobe). The second training phase taught AD and AE conditional discriminations; again, four three-member classes resulted (i.e., ADF classes, each including the name of a lobe of the brain, a brain schematic depicting the location of the lobe, and a deficit caused by damage to that lobe). Probes for class merger/expansion tested for the emergence of all possible relations between the B, C, D, and E stimuli in each class. To sum, the five-member classes documented by 80 different stimulus relations in this study were generated by teaching a subset of only 16, in lessons that required only about 15 minutes—a tremendous efficiency in teaching.

The applications of EBI to college coursework would seem to be limited only by the teacher's creativity. Indeed, a number of different subject matters have been targeted with this teaching approach. Statistical concepts have been a frequent choice for instructional programs (e.g., in addition to those above, Albright, Reeve, Reeve, & Kisamore, 2015; Fienup, Hamelin, Reyes-Giordano, & Falcomata, 2011). Other targeted content has included concepts from research design (e.g., Lovett, Rehfeldt, Garcia, & Dunning, 2011; Sella, Ribeiro, & White, 2014; Walker & Rehfeldt, 2012), from diagnostic categories of disabilities (Walker, Rehfeldt, & Ninness, 2010); and from classes of drug names in psychopharmacology (Zinn, Newland, & Ritchie, 2015). Although all of these studies have demonstrated tremendous teaching efficiency in the number of emergent match-to-sample relations generated, compared to the number trained, some have extended that efficiency significantly by including generalization tests of various kinds. Generalization testing strategies have included probes for topography-based responding that was not targeted in training (e.g., orally naming stimuli or writing answers; Albright et al., 2015; Sella et al., 2014); probes presented in a different format (e.g., on multiple-choice paper-and-pencil tests akin to those used in traditional classroom settings; Albright et al., 2015; Critchfield & Fienup, 2010); novel stimulus presentations (e.g., Albright et al., 2015; Walker & Rehfeldt, 2012); and follow-up maintenance tests (e.g., Albright et al., 2015; Walker & Rehfeldt, 2012).

Finally, in addition to evaluating the success of EBI with improvements at the level of the individual participant (e.g., from pretest to posttest), some studies have also compared performances across groups. As noted above, Fields et al. (2009) compared students who received training to those who did not, and found significantly stronger performances after training. Fienup and Critchfield (2011) compared the outcomes of EBI instruction to those from two different groups—one that received no training and one that received training for all possible relations (i.e., all of the relations that were trained as well as those that were probed for the EBI group). The students with no training performed poorly on the comparison tests; the EBI and full-training groups did equally well, but the EBI group had received significantly less training. Zinn et al. (2015) compared a group of students receiving EBI to each of two control groups trained on a randomly selected (rather than interrelated) subset of the EBI training and probe trials. One control group was matched to the EBI condition with respect to the accuracy criterion required for training; the other was matched with respect to the total number of training trials presented. Probe performances following equivalence training were significantly superior to those of the unstructured-instruction groups. Lovett et al. (2011) compared a group of students trained with EBI to a group that received a standard lecture on the same subject matter, presented as a video. The two groups performed equally well in this comparison. If EBI is to have widespread impact on university teaching practices, further comparisons across teaching formats will continue to be an important direction for future research.

	Set 1	Set 2	Set 3	Set 4
A	Frontal Lobe	Occipital Lobe	Parietal Lobe	Temporal Lobe
B	Function 1: Involved in movement	Function 1: Involved in perceiving visual information	Function 1: Involved in perceiving touch	Function 1: Involved in auditory processing
C	Function 2: Involved in higher cognitive functions	Function 2: Involved in integrating color and shape	Function 2: Involved in integrating sensations	Function 2: Involved in memory formation
D				
E	Damage causes impulsiveness	Damage causes problems with sight	Damage causes inability to perceive objects in space	Damage causes memory and hearing impairment

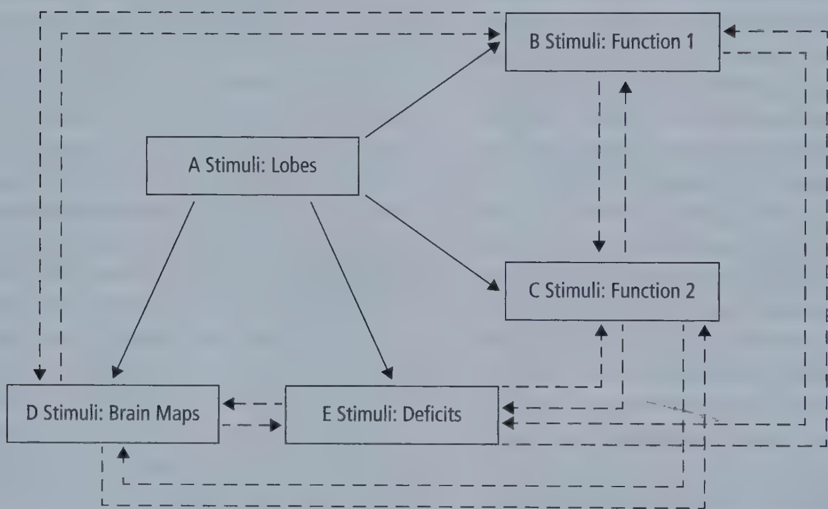


Figure 19.28 The top half of the figure presents the stimuli used in training; each row represents a stimulus set (A, B, C, D, E), and each column represents a potential equivalence class. The bottom half of the figure presents a schematic representation of all trained (solid arrows) and tested (dashed arrows) conditional discriminations.

From "Teaching Brain-behavior Relations Economically with Stimulus Equivalence Technology" by D. Fienup, D. Covey, and T. Critchfield, 2010, Reproduced with permission of John Wiley & Sons Inc.

APPLICATIONS STEMMING FROM ALTERNATIVE THEORETICAL APPROACHES TO RELATIONAL RESPONDING

Sidman's (1994, 2000) theoretical account of equivalence-class formation holds equivalence to be a basic behavioral process, akin to reinforcement or discrimination in that individuals do not have to learn how to behave in accordance with these processes. Just as no one had to teach you to respond more frequently when your behavior produces a reinforcer, in Sidman's view, no one had to teach you to produce the patterns of responding defined as equivalence. Rather, basic behavioral processes, like equivalence by this view, are a direct outcome of the reinforcement contingencies in place. In other words, arranging a three- or four-term reinforcement contingency will automatically result in equivalence performances as long as no methodological features constrain or interfere (e.g., by creating competing sources of

stimulus control), at least for most human participants. Further, Sidman's view is that all elements of the reinforcement contingency (e.g., for a four-term contingency, that would include the sample stimulus, the comparison stimulus, the response, and the consequence) can function as equivalence-class members in a given context (again, as long as no methodological issues interfere). In keeping with this emphasis on the reinforcement contingency as the basis for equivalence, the EBI literature reviewed thus far has highlighted many creative arrangements of three-, four-, and even five-term training contingencies, with focus on the make-up of the classes made possible by those contingencies.

Alternative theoretical accounts of equivalence performances have also been developed since Sidman's seminal work on the topic (e.g., Sidman, 1971; Sidman & Cresson, 1973; Sidman & Tailby, 1982), as the phenomenon has been explored and extended, both experimentally and conceptually. The two most prominent of these alternatives are naming theory, developed by Horne and Lowe (1996) and colleagues, and relational frame theory (RFT), developed by Hayes, Barnes, and colleagues (e.g., Hayes, Barnes-Holmes, & Roche, 2001). A complete review of the literature related to these accounts is well beyond the scope of this chapter; Chapter 20 provides a careful examination of work based on RFT, and the interested reader can consult the references on naming theory provided in this chapter. Some discussion of the two positions is warranted here, however, particularly because of the practical strategies for producing emergent (or derived) behavior that follow from each theory. Thus, the necessarily brief overviews that follow are intended to introduce a few representative illustrations of application work based on the naming and relational frame theories, to provide the reader with a basis for further study of these productive directions in applied behavior analysis.

One key theoretical feature shared by the naming theory and RFT distinguishes these accounts from Sidman's (e.g., (2000). In both cases, emergent or derived responding is seen as learned behavior, rather than as a direct outcome of a reinforcement contingency. In other words, a specific learning history is viewed as necessary before a given contingency could give rise to emergent performances. The nature of the training history held to be necessary differs across the two accounts, providing the basis for differences in the application approaches to be described. For both accounts, however, emphasis is placed on training that produces a generalized or **higher-order operant class**. Higher-order operants are defined in terms of general relations between antecedents and responses, rather than in terms of specific stimuli and responses, and any given instance of antecedent control is seen as just one example of a more general relation (e.g., Catania, 1998). Generalized imitation is considered a classic example of a higher-order operant (e.g., Catania, 1998; Dinsmoor, 1995). Imitation is typically viewed as a correspondence between the actions of a model and those of an observer, rather than in terms of any specific modeled action or response, and imitative repertoires are generalized when they include novel correspondences that have not been directly taught. To give a further feel for the central role played by this sort of conceptualization, manding (Catania, 1998; Skinner, 1957) and instruction following (e.g., Catania, 1998) stand as other examples of higher-order operant classes. The creation of higher-order operants requires teaching a number of examples, or multiple-exemplar training. For instance, a series of different modeled actions could be presented, with a child's responses approximating those of the model shaped and reinforced in the presence of each. After teaching a number of such correspondences, children have come to show appropriate approximations when presented with novel model actions (e.g., Baer, Peterson, & Sherman, 1967; Poulson, Kymissis, Reeve, Andreatos, & Reeve, 1991). Naming and relational framing are similarly viewed as higher-order operants, created by training many specific examples that ultimately provide for the acquisition of a more generalized class of discriminated responding.

Naming Theory

Naming has been suggested (Horne & Lowe, 1996) as a technical term for a higher-order operant in which an individual engages in bidirectional speaker and listener behavior with respect to a given object or event. In one example, a child might see a shoe, tact "shoe" (either overtly or covertly), hear herself say "shoe," and then exhibit listener behavior occasioned by hearing herself say "shoe"—perhaps by orienting to the shoe or placing her foot in it. The complete cycle is necessary to satisfy the requirements for naming in this technical sense, and naming is viewed as the basic unit of verbal behavior. To establish the naming operant, each component of the cycle (e.g., listener responses, tact responses) needs to be directly taught (e.g., Horne et al., 2004; Lowe, Horne, Harris, & Randle, 2002), and with many different stimuli, or exemplars (e.g., shoes, dolls, dogs). Once fully established, however, the entire cycle could be initiated from any point (seeing the shoe and then saying "shoe," or hearing "shoe" and then looking for the shoe, thus meeting the bidirectional dimension of the definition), and the acquisition of new instances of naming would require teaching only one component of the cycle (e.g., teaching a child to select a new toy given a novel name would be sufficient to produce the full naming relation). Figure 19.29 illustrates that seeing any number of different shoes could occasion saying "Shoe," and that hearing "shoe" could occasion listener behavior toward a number of different shoes, which would then occasion saying "Shoe," and so on. Naming is thus viewed as classifying, so that any two or more items or events with the same name would be treated similarly; they would occasion the same tacts and thus the same listener behaviors (e.g., Horne, Hughes, & Lowe, 2006; Lowe, Horne, & Hughes, 2005). In this way, naming is seen as responsible for categorization, equivalence performances, and various forms of transfer of function (Horne & Lowe, 1996).

Of course, a large number of the EBI programs described so far involved names (e.g., Sidman, 1971; Sidman & Cresson, 1973), whether presented as auditory samples, as written samples or comparisons, or as responses, and certainly the data are clear that providing stimulus names can facilitate equivalence-class formation (e.g., Randell & Remington, 1999, 2013; Sprinkle & Miguel, 2012). Although such findings are consistent with naming theory, they are also consistent with Sidman's theoretical position that any term in a reinforcement contingency, including a class-consistent response such as a

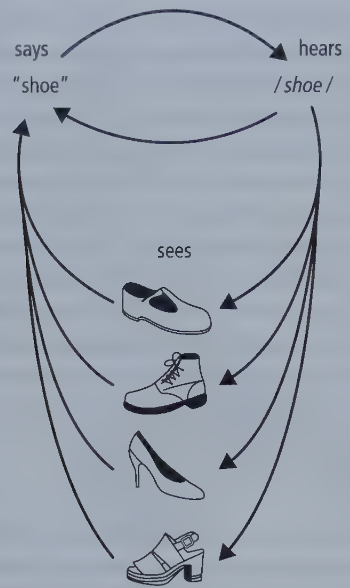


Figure 19.29 An illustration of the naming relation, where hearing "shoe" occasions listener behavior (e.g., orienting toward any of a number of different shoes), which in turn occasions tacting "shoe." The naming unit could be initiated from any point in the cycle.

From "On the Origins of Naming and Other Symbolic Behavior" by P.J. Horne and C.F. Lowe, 1996, Reproduced with permission of John Wiley & Sons Inc.

name, can become an equivalence-class member. Applications more unique to naming theory have sought to generate emergent performances with training procedures that target each of the individual components that constitute the complete naming unit. Indeed, well-developed protocols have been developed for teaching each of the listener and speaker skills involved (e.g., Horne & Lowe, 2000; Miguel & Petursdottir, 2009). Not surprisingly, the populations involved in these applications have been very young children or children with intellectual disabilities, for whom higher-order naming is not a part of the existing repertoire.

To illustrate, two studies compared the impact of listener training and tact training on the emergence of intraverbal responses involving a native language word and a second-language word (Petursdottir & Haflidadottir, 2009; Petursdottir, Olafsdottir, & Aradottir, 2008). A total of six typically developing 5-year-olds were involved. The stimuli included color photographs, native Icelandic words, and corresponding words from a different language (either Spanish or Italian). The children were able to name each picture in Icelandic at baseline, and training was conducted with two stimulus sets for each child. Two of the children received tact training only, two received listener training only, and two received each form of training for a different stimulus set. Tact training involved presenting a picture and teaching the child to produce its name orally in the second language. Listener training involved presenting three pictures and teaching the child to select the appropriate one when its second-language name was presented orally by the experimenter. Pretests and posttests presented each Icelandic word and asked the child to say its second-language name, and vice versa. Across the studies, both forms of training resulted in significant improvements from pretest to posttest, with tact training producing the greatest gains in emergent intraverbals (see also May, Downs, Marchant, & Dymond, 2016). Similar findings were obtained with a study of two 5-year-old children with autism (Grannan & Rehfeldt, 2012), in which pre- and post-intraverbal tests involved asking the child to produce the names of members of a category (e.g., body parts or vehicles). No correct responses were given on the pretests, but after listener and tact training, substantial improvements in intraverbal skills emerged.

Other studies have evaluated the impact of listener (Kobari-Wright & Miguel, 2014) and tact training (Miguel & Kobari-Wright, 2013) on categorization skills in young children with autism, a common target in many curricula designed for this population and one that frequently requires extended training. The training stimuli used here were nine cards picturing three examples of each of three categories of dogs (i.e., hound dogs, toy dogs, and working dogs). On each trial of the categorization test, a picture from a particular category was presented as a sample stimulus along with three additional pictures as comparisons, one from the same category and two from different categories. The categorization test was given twice before and twice following training; the second test of each pair also required the child to name the sample before the comparisons were presented. Accuracy was at approximately chance levels on the pretests, and training was then conducted with all nine cards. Each child received either listener training, in which she learned to select the correct picture from three comparisons when given its category name, or tact training, in which she was taught to produce the category name orally when shown a picture, or both if necessary. Of the four children who received listener training, three demonstrated near-perfect emergent performance on the category tests and also on posttraining tact tests. For the fourth, neither tact responses nor categorizations improved; when tact training was added, however, accuracy was perfect on the subsequent tests. Of the two children exposed to tact training first, both demonstrated accurate listener behavior and strong performance on the categorization tests, especially when a tact of the sample was also required on each test trial. Collectively, these data suggest that active tacting may be important for accurate categorization with this population, and that listener training alone can be insufficient for generation of tacting with some children. Similar results were obtained in a study with two young men with moderate intellectual disabilities and extremely limited vocal repertoires (Ribeiro, Elias, & Goyos, 2010) for whom signing was the targeted skill. During listener training, a video clip of each of six signs being performed was presented as the sample stimulus, along with three picture comparisons on each trial (AB training). After training, both men demonstrated emergent signing (tacting) and emergent manding with the same signs.

As a group, the studies described here illustrate an approach to producing novel performances that follows from the premises of naming theory. For certain participant populations, teaching critical components of the naming unit (i.e., listener behavior, tacting) can be effective in generating repertoires indicative of symbolic relations, classes, and categories.

Relational Frame Theory

In relational frame theory (RFT), framing is viewed as the key higher-order operant, and it is characterized as responding to one stimulus in terms of another (Hayes et al., 2001). RFT emphasizes that we learn many different types of relations in our lives (e.g., larger than, opposite of, different from), each of which can be considered a frame in that any pair of stimuli can function within it. Equivalence is thus just one of many frames (known as the *frame of coordination* in RFT terminology), although certainly an early and important one. As noted previously, framing is behavior and it must be taught; multiple-exemplar training is viewed as the basis for learning to frame. In a classic example, a mother might teach symmetry by engaging her child with a picture book. Mom could point to the picture of a dog, say "dog," and then ask, "Where's the dog?"—shaping, prompting, and reinforcing the child's pointing to the appropriate picture. This same interaction would unfold for many different pictures, with the child's symmetrical response directly trained in each case. Training with a number of exemplars allows for abstraction of the bidirectional relationship between names and pictures; this relation could then be applied to new examples without the need for further shaping, demonstrating emergent or derived relations, in RFT terms. To describe a performance as derived thus implies an active response on the part of the behaver—a relating or framing response that has been directly taught with many previously experienced exemplars. In the natural acquisition of many frames, early exemplars are likely to be characterized by real physical differences; for example, initial examples of "larger than" training are likely to involve pairs of stimuli in which one is physically greater in size than the other. Once abstraction of the "larger than" relation has been learned, however, it could be applied to novel stimulus pairs, including those for which the physical dimension of size is unimportant; for example, a learner could respond appropriately after being told that 7 is larger than 5, or that a dime is larger in value than a nickel.

Given the broad range of possible relations in our world, RFT introduces generic terms for the critical properties of framing. These include mutual entailment (e.g., if A is larger than B, then B is smaller than A; in equivalence relations, symmetry would be an example of mutual entailment); combinatorial entailment (e.g., if A is larger than B, and B is larger than C, then A is larger than C, and C is smaller than A; in equivalence relations, transitivity would be an example of combinatorial entailment); and transformation of function (e.g., if A is larger than B, and B is good, then A would be better; in equivalence relations, transfer of exactly the same function from one stimulus to another in its class would be an example). Also, given the number of relations possible, RFT emphasizes the critical role of contextual stimuli to indicate exactly which relation should be applied in any given instance. According to RFT, with the proper training history and given the appropriate contexts, frames can be combined in virtually limitless networks, providing for increasingly complex derived performances. Thus, framing is viewed as the basic behavioral unit not only of language but also of almost all complex cognitive functioning for humans. Indeed, at a conceptual level, RFT has been seen as relevant to domains ranging from the acquisition of basic relations (e.g., opposite, different) by young children to religion and spirituality (e.g., Barnes-Holmes, Hayes, & Gregg, 2001) to adult psychopathology and its treatment (especially in the form of Acceptance and Commitment Therapy; e.g., Hayes, Luoma, Bond, Masuda, & Lissis, 2006; Wilson, Hayes, Gregg, & Zettle, 2001).

In keeping with the breadth of the theory, applications of RFT have been varied with respect to targeted skills and participant populations. A few examples here must suffice for illustrating the focus on multiple-exemplar training and on emergent relational performances other than equivalence. (Because more than one relation is often targeted, RFT procedures can be a bit more complicated than many of those described thus far.) One early RFT application study (Berens & Hayes, 2007) asked whether multiple-exemplar training could establish a frame of comparison (more than and less than) with four typically developing 3- and 4-year-old girls who had low scores on pretests with this frame. For each test or training trial, either two or three pictures were presented (e.g., a smiley face, a lightning bolt, and a heart), and comparative relations between the stimuli were taught. For example, on a two-picture trial, the child might be told, "This (with a point to the smiley face) is more than that (with a point to the lightning bolt)." Next, the child was asked, "Which would you use to buy candy?" Probe tests were conducted with three different stimulus sets before training began and then after each phase of multiple-exemplar training. All training was conducted with a single stimulus set. The first training phase taught the "more than" relation with the same two pictures, and the picture designated as "more than" the other varied from trial to trial (i.e., A might be more than B on one trial, and B more than A on the next). Correct responses to the candy-buying question were reinforced across the repeated trials. The training thus provided multiple exemplars of the "more than" relation, and required control by the contextual stimulus ("more than") and the specifics of the instruction, as opposed to control by the physical properties of the actual stimuli involved. Subsequent training phases taught "less than" with two-stimulus trials, "more than" and "less than" with three-stimulus trials, and "more than" and "less than" in the same trial (e.g., "A is more than B, and C is less than B"). After the teaching steps were mastered, all four girls demonstrated emergent comparative relations with new stimulus sets. In many instances, accurate test performances were demonstrated on trial formats that hadn't yet been targeted in training. That is, performance on three-picture trials improved after training with two-picture trials, and performance on mixed-relation trials improved after the "more than" and "less than" relations were taught independently. In short, the multiple-exemplar training was successful in producing emergent performances, generalized across many different stimulus pairs and across test arrangements.

A similar approach was used to establish emergent manding for "more" or "less" in young children with autism (Murphy & Barnes-Holmes, 2009; Murphy, Barnes-Holmes, & Barnes-Holmes, 2005) and to generate derived tacting in a second language (Rosales, Rehfeldt, & Lovett, 2011). In the latter case, 3- to 4-year-old Spanish speakers were taught to select the appropriate picture when presented with its English name (i.e., listener training). When this failed to produce tacting in

English, additional listener training with multiple novel words was followed by tact training with the same words. This resulted in substantial improvements in emergent tacting with the original words after listener-training only, with evidence of maintenance for many of the tacts after 1 month. Results like these demonstrate the value of multiple-exemplar approaches for language training programs, where generalized or higher-order responding is especially important. Other interesting work with multiple-exemplar training has sought to establish generalized perspective-taking skills in both typically developing children and those with autism (e.g., Barnes-Holmes, McHugh, & Barnes-Holmes, 2004; Heagle & Rehfeldt, 2006). After teaching multiple instances of each relation, children have shown appropriate derived responses with new examples for relations of "I" as contrasted with "You," "Here" as contrasted with "There," and "Now" as contrasted with "Then" (Heagle & Rehfeldt, 2006), all thought to be critical components of the perspective taking that is necessary for more complex reasoning and for social skills (Barnes-Holmes, Hayes, Dymond, & O'Hara, 2001). Indeed, intensive multiple-exemplar training for four frames (same, opposite, more, and less) was reported to correlate with increased Wechsler Intelligence Scale for Children (4th ed.; WISC IV) scores (a common measure of intelligence quotient [IQ]) for fifteen 11- to 12-year-old typically developing children, and with increased scores on the Differential Aptitude Test (5th ed.) for 30 typically developing 15- to 17-year-olds (Cassidy, Roche, Colbert, Steward, & Grey, 2016), an intriguing outcome certainly worthy of further study.

With respect to adult participants, multiple-exemplar training approaches have been used successfully to establish rather complex mathematical skills (i.e., matching standard and factored forms of mathematical formulas with each other and with graphical representations of the formulas) in college students (e.g., Ninness et al., 2005), where the role of contextual stimuli (Ninness et al., 2006) and of multiple frames (same as and reciprocal of; Ninness et al., 2009) has also been demonstrated. Certainly these results add to those of studies on equivalence-based instruction for efficient approaches to teaching college-level academic content.

In another program of study, multiple-exemplar training has been utilized to impact the behavior of recreational and pathological gamblers (e.g., Dixon & Holton, 2009; Hoon, Dymond, Jackson, & Dixon, 2007, 2008; Zlomke & Dixon, 2006). To illustrate, five pathological gamblers completed delay discounting tests in which a series of choices were made between an amount of money to be delivered now and an amount to be delivered at some future time (e.g., a choice between $500 today and $700 in 12 weeks). Past studies have established strong reliability for the discounting task (e.g., Lagorio & Madden, 2005) and have shown that delayed consequences hold less value for gamblers than for nongamblers; that is, gamblers show a greater degree of discounting (e.g., Dixon, Jacobs, & Sanders, 2006). In the present study, the discounting task was presented with the smaller, sooner choice always shown in purple and the larger, later choice shown in pink. Multiple-exemplar training was then conducted in a match-to-sample format as follows. On each training trial, either a pink or a purple square was presented as a contextual stimulus, followed by a sample stimulus and three comparisons from the A stimulus set illustrated in Figure 19.30. When the purple context was presented, selecting the comparison

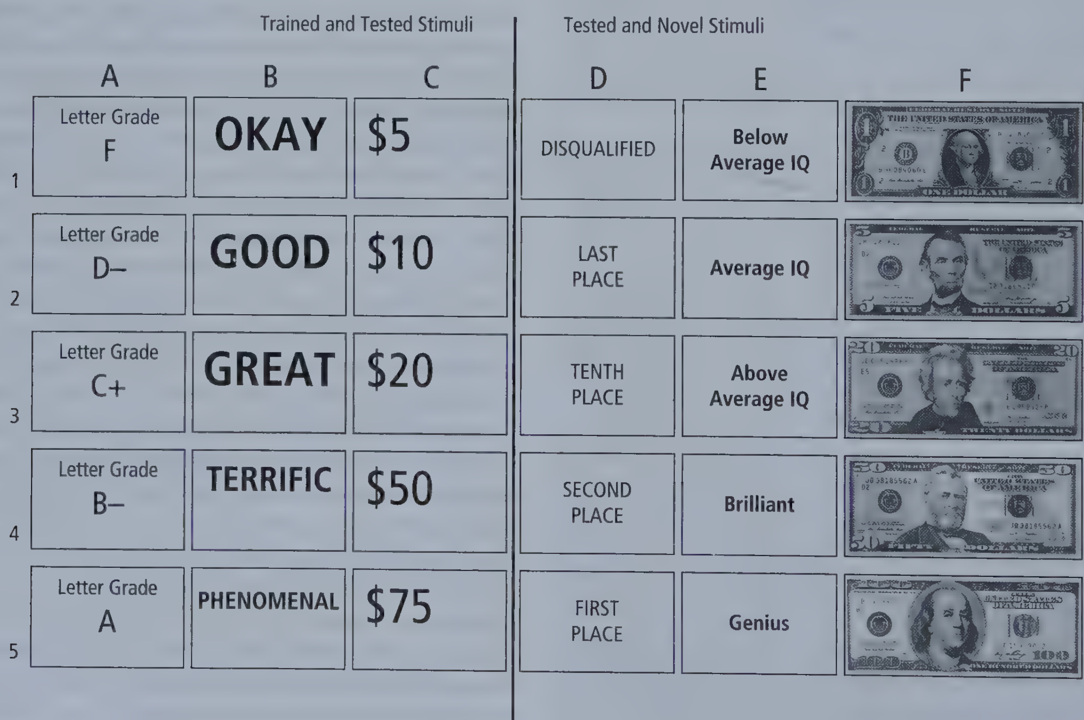


Figure 19.30 Stimulus sets used in multiple-exemplar conditional discrimination training.
From "Altering the Magnitude of Delay Discounting by Pathological Gamblers" by M.R. Dixon and B. Holton, 2009, Reproduced with permission of John Wiley & Sons Inc.

that was worse than the sample produced a reinforcer (e.g., given C+ as the sample, selecting D− rather than B− or A); when the pink context was presented, selecting the comparison that was better than the sample produced a reinforcer (e.g., given C− as the sample, selecting B− rather than D− or F). Next, training trials were conducted with stimuli from Sets A, B, and C (i.e., multiple-exemplar training), and then trials from Sets A to F were presented with no consequences, to assess control by the contextual stimuli with novel stimuli. Test performances confirmed that the purple context controlled selections of the "worse than" comparison, while the pink context controlled selections of the "better than" choice. When the discounting task was repeated as a posttest, all five participants demonstrated a substantial shift in their discounting, such that the smaller, sooner choices shown in purple were selected far less often than on the pretest. In short, the reliable discounting outcome was changed dramatically by simple stimulus cues, the function of which had been established via a brief multiple-exemplar teaching program with "better" and "worse" relational frames. These results are worthy of follow-up given their implications for the development of assessments as well as possible interventions with this clinical population.

To sum, the studies described in this section emphasize an approach to generating novel symbolic behavior that reflects the tenets of relational frame theory. Ample evidence shows that multiple-exemplar training can help to establish the sorts of generalized or higher-order forms of responding that contribute to complex symbolic skills.

CONCLUDING REMARKS

This chapter has presented a case for equivalence-based instruction as one of the most exciting directions in research and practice in applied behavior analysis. Indeed, the powerful generativity of an equivalence approach, demonstrated repeatedly across participant populations, procedures, settings, and stimulus sets, demands the attention of practitioners, applied scientists, basic laboratory scientists, and conceptual theorists alike, and ensures the sort of ongoing translational effort that characterizes the best of behavior analysis. The pattern of work to date suggests that ever more effective and efficient equivalence protocols will continue to emerge, as scientists learn to combine critical features of the most successful programs. There can be little doubt that the studies described and reviewed in this chapter represent only the tip of the proverbial iceberg in terms of what might be accomplished with EBI approaches.

SUMMARY

Research Foundations and Core Concepts

1. Stimulus equivalence describes a behavior analytic approach to understanding and establishing symbolic function. Our field's interest in the equivalence phenomenon can be traced back to early work by Sidman (1971; Sidman & Cresson, 1973) on reading comprehension.

2. Arbitrary match-to-sample procedures are commonly used in equivalence studies to teach the necessary baseline performances and to test for emergent (or untrained) ones. Arbitrary match-to-sample training involves a discrete-trial four-term contingency arrangement with physically dissimilar stimuli. A trial begins with the presentation of a sample stimulus (e.g., A1 or A2); an observing response to the sample produces two or more comparison stimuli (e.g., B1 and B2). Selection of the comparison designated as correct for that sample results in a reinforcer, while selection of any other comparison does not. Arbitrary match-to-sample training produces conditional discriminations, in which discrimination between the comparison stimuli is conditional on the sample presented (e.g., given an A1 sample, choosing B1; given an A2 sample, choosing B2).

3. After teaching two interrelated conditional discriminations (e.g., AB and BC), probe trials are presented to test for emergence of the untrained relations that define equivalence. There are no programmed consequences for probe trials; thus, consistent patterns of responding on these trial types are described as emergent or derived. The critical probe types include:
   - *reflexivity tests*, that is, tests for generalized identity matching with all training stimuli, or AA, BB, and CC; in mathematical set theory, reflexivity is defined as A = A.
   - *symmetry tests* for the reversibility of sample and comparison functions, or BA and CB; in mathematical set theory, symmetry is defined as, If A = B, then B = A.
   - *transitivity tests* for relations between stimuli that were never presented together during training but that share a trained relation with a nodal stimulus, or AC; in mathematical set theory, transitivity is defined as, If A = B, and B = C, then A = C.
   - *combined or equivalence tests* in which both symmetry and transitivity are evaluated simultaneously, or CA.

4. Positive outcomes on all probe types document interchangeability among a given group of stimuli and allow the conclusion that an equivalence class has formed. That is, responses on reflexivity, symmetry, and transitivity or equivalence trials show that any member of the class can serve effectively as sample or as comparison with any other class member. The equivalence class consists of all such interchangeable stimuli (e.g., A, B, and C, in this example).

5. The equivalence outcome is significant because:
   - Many stimulus relations are generated after teaching only a small number (i.e., there is an impressive "bang

for your buck" effect). This generativity carries tremendous implications for behavioral applications.

- The defining criteria for equivalence provide a compelling operational definition of symbolic function, and make possible a behavioral analysis of a range of complex human performances (e.g., categorization, semantic correspondence, creativity).

6. The literature on stimulus equivalence is characterized by a true translational approach. Advances in our understanding and application of equivalence have resulted from a necessary and continuing interplay between basic laboratory research, applied behavior analysis, and conceptual developments.

7. Several key outcomes of equivalence approaches have contributed to the effectiveness of EBI programs in important ways. These include:
   - ***class formation***—typically measured as the percentage of probe trials on which a comparison selection is consistent with equivalence, for each defining probe type. Such measures are often compared in a pretest, train, posttest design, or in a multiple baseline across stimulus sets experimental design.
   - ***delayed emergence***—a relatively common observation in which equivalence performances grow stronger with repeated testing.
   - ***class expansion***—increasing the size of an established equivalence class to include an additional member or members. This is typically accomplished by teaching a new relation between a novel stimulus and an established class member. Class expansion is verified by emergent relations between the novel stimulus and the rest of the established class members.
   - ***class merger***—a combining of independent equivalence classes such that a larger class is established. This is typically accomplished by teaching a new relation between one member of each of the classes to be combined. Class merger is verified by emergent relations between the other stimuli from the formerly independent classes.
   - ***transfer of function***—the finding that teaching a new function for one member of an established equivalence class results in the same function holding for all members of the class.
   - ***contextual control***—stimulus control requiring three levels of antecedent stimuli, such that the functions of the stimuli in a conditional discrimination vary depending on the context. Contextual control training requires a five-term contingency. It allows for the same stimuli to be members of more than one equivalence class, depending on the context.

Designing Equivalence-based Instruction

8. A number of procedural variables have proven important for the successful generation of equivalence classes, and for prevention of problematic stimulus-control patterns that could compete with demonstrations of class formation. These variables are considered in the context of making decisions about training and testing procedures.

9. In planning match-to-sample training procedures, critical decisions will include the number and type of stimuli to be used, the mode of stimulus presentation, inclusion and type of observing response, content of the instructions to be presented, the approach to balanced trial presentations, the necessity of special training steps, the type of training structure to be used, the mastery criteria for training steps, and the arrangement of consequences for training trials.

10. In planning the testing protocol, critical decisions will include the composition of the probe-trial blocks, the arrangement of consequences within the probe-trial blocks, the order of probe-type presentation, and the criteria to be used for judging class formation.

11. In addition to the standard match-to-sample training and testing approach to equivalence, several effective procedural variations have been identified that may carry benefits for particular populations or for certain EBI targets. These include:
    - ***stimulus compounds***—when a two-part stimulus (e.g., a word presented together with a picture) is presented as the sample in a MTS procedure, for example, each element of the compound can come to function as an independent equivalence-class member, thus increasing the size of the class created by a given amount of training.
    - ***class-specific reinforcement***—providing a unique reinforcer for discriminated responses as a function of the antecedent's intended class (e.g., in match-to-sample training, choosing B1 given A1 would produce Reinforcer 1, while choosing B2 given A2 would produce Reinforcer 2). Class-specific reinforcers can function as members of an equivalence class, again increasing the size of the class made possible by a given amount of training.
    - ***three-term contingency training***—simple discrimination training is mastered more easily than conditional discrimination training by some learners, and can produce equivalence-class formation.

Applications and Generality

12. The generality of the equivalence outcome has been repeatedly confirmed and extended by creative applications of EBI across populations, settings, and instructional targets.

13. Work with children and adults with intellectual disabilities, as well as young typically developing children, represents the largest category of EBI applications. Targets have included language- and number-skills instruction, as well as other types of functional skills (e.g., music, academic content, Braille reading).

14. Important work with adult clinical populations includes examples with health-related behaviors and skills impacted by brain injury; investigations in the field of gerontology are also being explored.

15. Equivalence-based instruction targeting university course content has had promising results across a range of subject matters and modes of presentation.

Applications Stemming from Alternative Theoretical Approaches to Relational Responding

16. Naming theory and relational frame theory (RFT) provide differing accounts of the basis for novel relational responding. Both emphasize that novel responding is due to a special learning history, rather than a direct outcome of a reinforcement contingency (as in Sidman's view). Naming Theory and RFT differ, however, with respect to the nature of the learning history that is required. These differences give rise to unique applications for generating emergent or derived responding.

17. For naming theory, naming is a technical term for a bidirectional higher-order unit of speaker and listener behavior that is seen as responsible for any form of classifying, including equivalence performances. Following from this view, applications designed to promote equivalence include instruction in the components of naming—listener behavior and tacting. For participants lacking these skills, targeting them directly can be effective in producing response patterns indicative of classes and categories.

18. For RFT, framing, or responding to one stimulus in terms of another, is the key higher-order operant. It is established by teaching many different examples of a given relation, which allows for abstraction and then application of the relation to new examples, given the proper context. RFT emphasizes that equivalence is just one example of the many relations we learn in this manner. Following from this view, applications designed to promote novel, or derived, relational responding feature multiple-exemplar training as the key instructional approach. A range of examples has provided evidence for the production of generalized response patterns that can contribute to complex symbolic skills.

KEY TERMS

class expansion
class merger
class-specific reinforcement
conditional discrimination
contextual control
derived stimulus relations
emergent stimulus relations
equivalence-class formation
equivalence test
exclusion
higher-order operant class
matching-to-sample procedure
nodal stimulus (node)
reflexivity
simple discrimination
simple-to-complex testing protocol
stimulus-control topographies
stimulus equivalence
symmetry
training structure
transfer of function
transitivity

MULTIPLE-CHOICE QUESTIONS

1. Interchangeability of function can be assessed by
   - **a.** Reflexivity
   - **b.** Symmetry
   - **c.** Transitivity
   - **d.** All of the above

   Hint: (See "Transfer of Function")

2. In simple-to-complex testing protocol, probe types are introduced sequentially
   - **a.** Beginning with symmetry followed by transitivity
   - **b.** Beginning with transitivity followed by symmetry
   - **c.** Beginning with reflexivity followed by transitivity
   - **d.** Beginning with reflexivity followed by symmetry

   Hint: (See "Test Order")

3. Which of them is a critical property of framing as proposed in relational frame theory?
   - **a.** Mutual entailment
   - **b.** Transformation of function
   - **c.** Combinational entailment
   - **d.** All of the above

   Hint: (See "Relational Frame Theory")

4. In the study by Sidman, 1971; Sidman & Cresson, 1973, three member equivalence classes (________ , ________, written word), can be increased to include a potentially ________ number of members based on class ________and class ________.
   - **a.** Spoken word, picture, unlimited, expansion and merger
   - **b.** Touch, sound, unlimited, contraction and demerger

c. Touch, sound, limited, contraction and merger

d. Spoken word, picture, limited, expansion and demerger

Hint: (See "Class Expansion and Class Merger")

5. In the EBI program designed by Fineup, Covey and Critchfield (2010), to document five-member classes, how many subsets were used?

a. 16 subsets produced 80 different stimulus relations in 15 minutes

b. 80 subsets produced 16 different stimulus relations in 15 minutes

c. 80 subsets produced 15 different stimulus relations in 25 minutes

d. 15 subsets produced 80 different stimulus relations in 25 minutes

Hint: (See "In the College Classroom")

6. Methods used for presenting probe trials is/are:

a. Massed Testing

b. Interspersing

c. Both of the above

d. None of the above

Hint: (See "Composition of Probe-trials Blocks")

7. Component Simple Discriminations training approach is based on the recognition that accurate conditional discrimination performances require:

a. Conditional control by the sample

b. Successive discrimination between the sample stimuli

c. Simultaneous discrimination between the comparison stimuli

d. All of the above

Hint: (See "Instructional Programming")

ESSAY-TYPE QUESTIONS

1. Describe the training structures used in equivalence-based instructions.

   Hint: (See "Training Structures")

2. Define higher-order operants with an example.

   Hint: (See "Applications Stemming from Alternative Theoretical Approaches to Relational Responding")

3. Discuss transfer of function. Elaborate on the negative impact of equivalence-class formation on an individual's functioning.

   Hint: (See "Transfer of Function")

NOTES

1. As we will see in the section on Relational Frame Theory, equating these terms can be problematic, because the latter carries important theoretical implications about the basis for responding. Thus, "emergent performances" is often preferred as a theory-neutral description.
2. Exemplary experimental analyses of the complex behavior patterns mentioned have been conducted by Augustson and Dougher, 1997; Augustson, Dougher, and Markham, 2000; Challies, Hunt, Garry, and Harper, 2011; Dougher, Augustson, Markham, Wulfert, and Greenway, 1994; Griffee and Dougher, 2002; Guinther and Dougher, 2010; Hayes, Thompson, and Hayes, 1989; Kohlenberg, Hayes, and Hayes, 1991; Moxon, Keenan, and Hine, 1993; and Watt, Keenan, Barnes, and Cairns, 1991.
3. EBI typically targets at least three classes. The example here is simplified for illustrative purposes.
4. With many populations, probes for reflexivity are omitted, primarily because identity matching can be assumed and is likely the result of experiences prior to the EBI training.

CHAPTER 20

Engineering Emergent Learning with Nonequivalence Relations

Thomas S. Critchfield and Ruth Anne Rehfeldt

LEARNING OBJECTIVES

- Understand what nonequivalence relations are and why do they matter.
- Discuss the vocabulary of nonequivalence relations.
- Explain some types of nonequivalence relations.
- Discuss the theoretical foundations of nonequivalence relations.
- Understand the connections between nonequivalence relations and psychological constructs: knowledge structures, self-concept, and intellectual development.
- Understand the connection between derived stimulus relations and general well-being.

Chapter 19 explained how equivalence ("sameness") relations[1] function and how they can be harnessed to generate valuable new learning. But equivalence is only one beast in a menagerie of ways in which stimuli can be symbolically and behaviorally linked. **Nonequivalence relations**—those in which stimuli are related on some basis other than "sameness"—are a big part of how people make sense of, and function effectively in, the world around them. We believe, therefore, that every behavior analyst needs a solid working understanding, and mastery, of both the general concept of stimulus relations and the specific domain of nonequivalence relations (e.g., Critchfield, 2018; Critchfield, Barnes-Holmes, & Dougher, 2018).

Our discussion of nonequivalence relations is intended to complement Chapter 19's discussion of equivalence relations. To properly understand nonequivalence relations, you need grounding in stimulus relations generally. By mastering the conceptual foundations of equivalence relations and the procedures that have been used to establish them, you prepare yourself to benefit from the present chapter.

This chapter is divided into six sections. The first section explains generally what nonequivalence relations are and why they matter. The second section presents basic vocabulary for talking about nonequivalence relations and phenomena associated with them. This is necessary because the vocabulary of equivalence relations is not always adequate for describing nonequivalence relations. The third section describes selected types of nonequivalence relations and illustrates their everyday importance. The fourth section introduces key theoretical concepts relevant to nonequivalence relations. The fifth section discusses the role of nonequivalence relations in big-picture psychological constructs that often are of interest in the practical world. The final section explains how derived stimulus relations, including nonequivalence relations, provide a basis for enhancing general well-being in ways that are quite different from what is typical in applied behavior analysis (ABA).

Although this chapter highlights many opportunities to incorporate nonequivalence relations into applied interventions, we will say little about practical tactics of implementation, for two reasons. First, because the behavioral repertoires of interest can be complex, procedures for creating them often are as well. Before developing protocols to teach nonequivalence relations, the interested student may need to digest other sources that provide the required detail (e.g., Rehfeldt & Barnes-Holmes, 2009). Doing so, however, requires conceptual skills that develop with practice. Thinking about derived stimulus relations is a detail-heavy logical process that does not come naturally to many people, and the present chapter is an introduction to this process. Second, in many respects the mechanics of building equivalence relations, as discussed in Chapter 19, also apply to the construction of

➠The authors thank Bryan Roche, Dan Fienup, and Daniel Assaz for providing helpful feedback on a draft of this chapter.

nonequivalence relations (e.g., Luciano et al., 2009). For example, both depend on directly teaching baseline relations consisting of overlapping conditional discriminations. In both cases, baseline relations must be learned before untaught relations can arise, and two stimulus classes with a shared member can merge to form a larger class—and so forth. Chapter 19 is a guide to these basics, and the present chapter builds on it by expanding the range of derived stimulus relations that can be the focus of intervention.

WHAT ARE NONEQUIVALENCE RELATIONS? WHY DO THEY MATTER?

To explain nonequivalence relations, we'll begin by expanding on a general point from Chapter 19: Most stimulus relations are **arbitrary relations** (Hayes, Fox, et al., 2001), stimuli that "go together" not because they are physically identical or because a law of the universe demands this, but rather because social-verbal reinforcement contingencies teach people to relate them in a certain way. Building on an example presented in Figure 19.7, there is no inherent reason that a token exchangeable for goods and services has to be called a *dime*, to be a disk of silver-colored metal, or to have its value represented with "10¢." Humans have arbitrarily created these symbols and the relations among them (such as the equivalence class in Figure 19.7). Had the token been a lump of pink salt and been called *armadillo*, for instance, the equivalence class in Figure 19.7 would look different but function the same.

Because relations among stimuli are arbitrary, the variety of possible relations is limited only by the repertoires and experiences of humans. People create all kinds of symbols and devise many different ways to compare and contrast them. How they do this is a core feature of how they understand the world, as each bit of the world is given meaning through the other bits of the world to which a person's experiences connect it. Consider the word *beer*, which is an arbitrary printed symbol for a particular beverage. The behavioral functions of *beer* are many and nuanced. To get a sense of this, think about what *beer* means to you by answering the questions below (adapted from Barnes-Holmes, Hayes, Dymond, & O'Hora, 2001):

- What are other names for beer?
- What are the attributes of beer?
- What is beer composed of?
- What is beer similar to?
- Where is beer found?
- What is beer different from?
- What is beer an example of?
- What does beer include?
- What is beer better than?
- What does beer do?
- What came before beer?

Notice that all of these questions demand an active response; how you understand *beer* is evident in how you respond to it. Notice, too, that all of the required responses involve relating the target word to something else. The behavioral functions of *beer* are inherent in the things to which people relate it.

Figure 20.1 provides a few sample answers to the *beer* questions. The resulting diagram is reminiscent of the cognitive psychology concept of a *semantic network*, a hypothetical construct intended to model how bits of knowledge are organized in long-term memory.[2] Based on studies using a variety of methods (such as free association, in which people are presented with a word and asked for the first word that comes to mind), many cognitive researchers have concluded that bits of stored knowledge connect to one another through shared meaning (e.g., Collins & Loftus, 1975).

The semantic network construct is worth mentioning partly because it predates the earliest research by behavior analysts on derived stimulus relations (Sidman, 1971), and thus psychologists outside of behavior analysis tend to be familiar with it. Also, despite the fact that behavioral psychology and cognitive psychology are often portrayed as intellectual archenemies (Uttal, 2000), the semantic network construct incorporates at least three observations with which behavior analysts agree (see Hayes, Barnes-Holmes, & Roche, 2001; Reese, 1968; Sidman, 1994). First, as Figure 20.1 suggests, people relate various "bits of the world" in a *lot* of different ways. Second, the relationships of interest are forged through experience, such that each individual will have a relational network as unique as her learning history. Third, the relationships of interest are, in essence, the fundamental units of *meaning*, because no "bit of the world" is understood separately from its connection to others. A person with no *beer* experiences cannot answer any of the questions listed above. That person does not know the meaning of *beer*.

Behavior analysts believe they have improved on the semantic network concept in two ways. First, cognitive psychologists assume that "bits of knowledge" and the relationships among them are somehow represented (stored) inside the brain (Collins & Loftus, 1975). Behavior analysts say instead that each of us learns various ways of responding to stimuli like *beer*, and two stimuli are related if the way we respond to them is in some way intertwined (Critchfield et al., 2018; Reese, 1968). This is a network of behavior, not of knowledge.[3] Second, although most observers agree that relational networks arise from individual experience, cognitive psychologists have devoted more effort to describing the characteristics of already-formed networks than to understanding their formative experiences. By contrast, behavior analysts have typically tried to understand relational networks by studying their acquisition (Zentall, Galizio, & Critchfield, 2002). Because applied interventions exist to engineer changes in behavior, this approach, which focuses on experiences that create relational networks, is more informative than one that seeks to describe the features of relational networks that are the product of unspecified prior learning (Critchfield & Fienup, 2008).

As Chapter 19 explained, stimuli in the relational networks called equivalence classes are united by "same as" relations. For instance, *beer* has several synonyms (e.g., *cold one*, *brewski*, and *draft*) and together these stimuli may form an equivalence class. As Figure 20.1 indicated, however, *beer* can be important in many ways other than what it *is*. In fact, one can fully understand *beer* only by knowing what it is not, what it is similar (but not identical) to, what it's an exemplar of, what kinds of beer exist, and so on. This is the essence of nonequivalence relations.

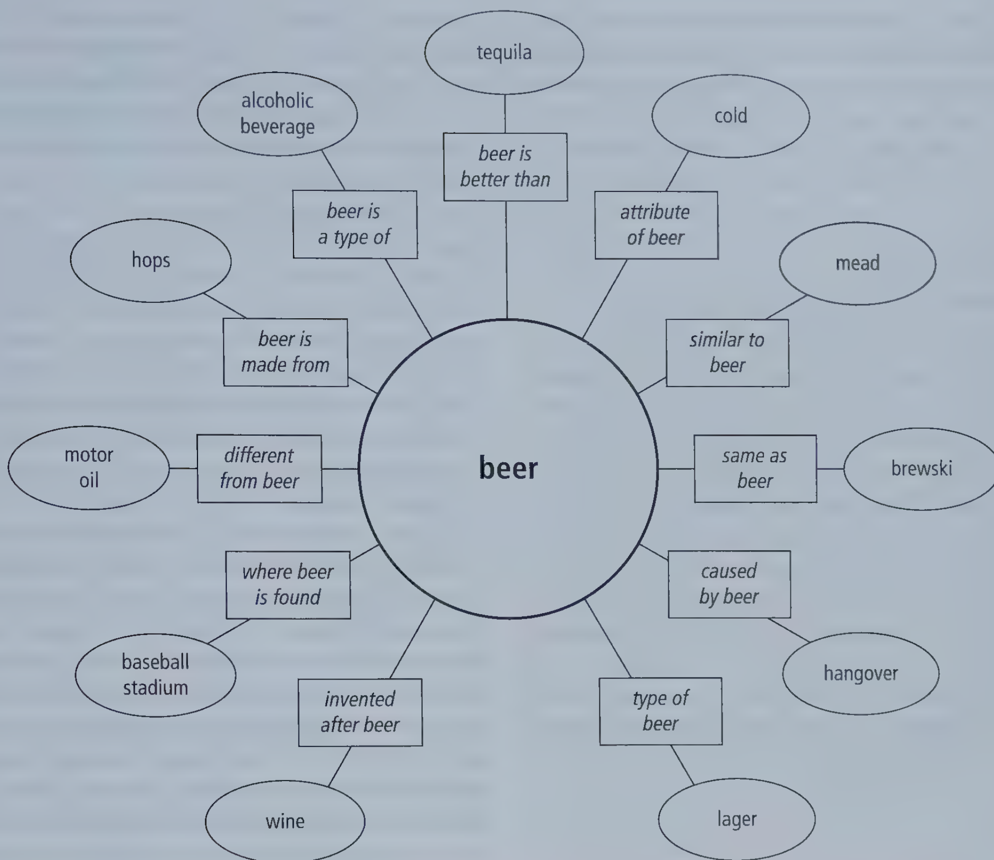


Figure 20.1 A variety of things to which *beer* is related, illustrating the wide variety of stimulus relations. Except for the "same as" example, all are nonequivalence relations.

In distinguishing nonequivalence relations from equivalence relations, it is important to note that the former sometimes have been included in behavioral research without proper acknowledgment. Figure 20.2 illustrates by reference to a study in which college students learned neuroanatomy concepts (Fienup, Covey, & Critchfield, 2010). As part of the lesson, students saw the term *frontal lobe* as sample stimulus and, from several pictures of brains, learned to select the one on which the frontal lobe was indicated by an arrow. With *frontal lobe* as the sample stimulus, the students also learned to select *impulsiveness* from among several alternatives as a symptom caused by damage to this lobe. The lesson was described as *equivalence-based* instruction, but it did not really involve "same as" relations. Rather, it involved a spatial relation (where the brain structure called *frontal lobe* is located) and a causal relation (what symptom is produced by damage to the brain structure called *frontal lobe*).

In one sense, it doesn't matter what Fienup and colleagues called the stimulus relations in Figure 20.2, because their students both mastered what was directly taught and demonstrated untaught relations (e.g., when shown the location of the frontal lobe, they could indicate that *impulsiveness* was caused by damage to that area). This was possible because the match-to-sample procedures used in the study gave students limited options for how to respond, so they chose comparison stimuli that were related to the sample in *some* fashion.[4] Difficulties might have arisen in a different kind of test. For instance, imagine that the students were shown a picture of the frontal lobe and asked to say its name. A student who answered *impulsiveness* (treating the brain picture, *frontal lobe*, and *impulsiveness* as "the same") would get the question wrong. What a person "needs to know," therefore, depends on

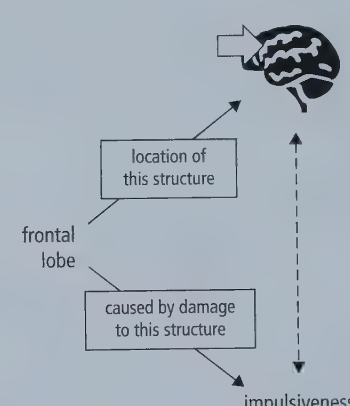


Figure 20.2 A neuroanatomy stimulus class that was established in college students. The published report called this an equivalence class, but it is more accurately described as a nonequivalence class.

Based on Figures 1 and 4 of "Teaching Brain-Behavior Relations Economically with Stimulus Equivalence Technology," by D. M. Fienup, D. P. Covey, and T. S. Critchfield, 2010, *Journal of Applied Behavior Analysis*, *43*, 19–33.

contingencies of reinforcement that often involve nonequivalence relations (for more on this, see the "Theoretical Foundations" section). In fact, nonequivalence relations appear to be at the root of many abilities thought of as representing "deep understanding," "higher thought," or "complex cognition" (Hayes et al., 2001), a point on which we will comment later.

THE VOCABULARY OF NONEQUIVALENCE RELATIONS

Much of the vocabulary coined to describe equivalence relations is too narrow to describe nonequivalence relations, so more inclusive alternatives have been proposed (Hayes, Fox, et al., 2001). Recall from Chapter 19 that stimulus equivalence classes are defined by relations of "sameness." Two emergent manifestations of "sameness," *symmetry* and *transitivity* (Sidman, 1994; see also Chapter 19),[5] are illustrated in the top left panel of Figure 20.3 as part of how three words (*overdue*, *tardy*, and *late*) come to "mean the same thing" and are treated by a learner as interchangeable. By *emergent* we mean relational responding that experience does not directly teach, but that arises as an indirect function of experience (Chapter 18 similarly referred to *generative* effects). Emergent relations also are called **derived relations**, and this is the term we will use moving forward.

Entailment

Symmetry is when, after learning to respond to stimulus B as if it were stimulus A, a learner spontaneously can do the reverse. Symmetry implies mutual "sameness." For example, in the top left panel of Figure 20.3, a learner who is shown *overdue* and taught to identify *tardy* as "the same" should also be able to identify *overdue* as "the same" when given *tardy*. Now consider the nonequivalence relation shown in the top right panel of Figure 20.3. A learner is shown *early* and taught to identify *tardy* as "opposite." The reverse of this directly taught relation is expected to emerge spontaneously: Shown *tardy*, the learner should be able to spontaneously identify *early* as "opposite." Because not all derived relations involve "same," the reverse of a learned relation can be of a different type than the learned relation. For instance, if a learned relation is that *25¢* is "more than" *10¢*, the reverse relation will be that *10¢* is "less than" *25¢*. The more general term **mutual entailment** applies to any derived bidirectional relation in which one "direction" has been directly learned (Hayes, Fox, et al., 2001).[6] Symmetry is a special case of mutual entailment, but not all mutual entailment is symmetry.

Transitivity is when a learner can spontaneously treat two stimuli as interchangeable ("same") because they are related indirectly through "sameness" with a common third stimulus. In the top left panel of Figure 20.3, a learner who is shown *tardy* should be able to spontaneously identify *late* as "the same," and vice versa. Now consider the nonequivalence relation in the top right panel of Figure 20.3. The derived relation between *tardy* and *late* ("same") is different from that between these words and their common associate, *early* ("opposite"). *Tardy* and *late* qualify as "same"—but not because they are both "the same" as a common third stimulus. Rather, in a case of "the enemy of my enemy is my friend," *tardy* and *late* both are opposites of the same word. The general term *combinatorial entailment* applies to any derived relation involving stimuli that are "linked"

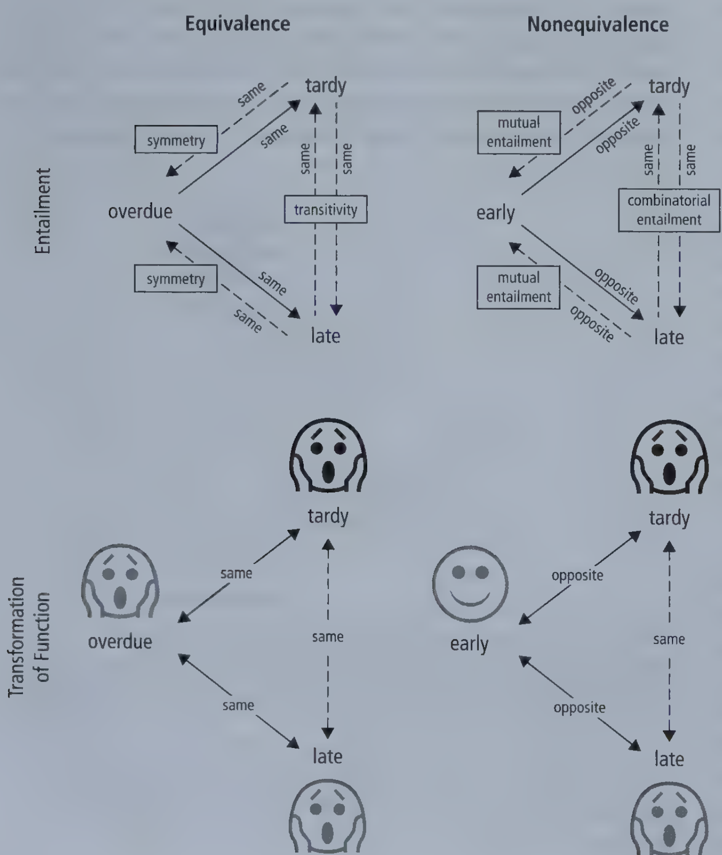


Figure 20.3 Relationships between terms used to describe derived relations in equivalence and nonequivalence stimulus classes. Top: Mutual entailment and combinatorial entailment as replacements for symmetry and transitivity, respectively. Bottom: Transformation of function as replacement for transfer of function. Emoticons represent emotion functions (fear and happiness); those in black are acquired through direct experience and those in gray are the result of transformation of function.

via common associates (Hayes, Fox, et al., 2001). To be more precise, **combinatorial entailment** is the relation involving two stimuli that both participate in mutual entailment with some common third stimulus. Transitivity is a special case of combinatorial entailment, but not all combinatorial entailment is transitivity.

Transformation of Function

Another defining feature of equivalence classes, as described in Chapter 19, is *transfer of function*. Chapter 19 noted that equivalent stimuli come to function in the same manner (control the same responses), and this applies to any new behavioral function acquired by any of the stimuli in a class. For instance, Dougher and colleagues showed that when classical conditioning made one stimulus in an equivalence class fear-eliciting, all of the other stimuli in the class became fear-eliciting as well (Dougher, Augustson, Markham, Greenway, & Wulfert, 1994; see also Dymond, Schlund, Roche, & Whelan, 2014). The bottom left panel of Figure 20.3 illustrates this effect. Imagine that a sensitive student is one day harshly reprimanded for being *tardy* to class. This scares her and makes the word *tardy* a fear-eliciting stimulus. If *tardy* is part of the equivalence class shown in Figure 20.3, then words like *overdue* and *late* will also come to elicit fear. "Transfer" of function in this case is the propagation of a single behavioral function across stimuli that are "the same" (an equivalence class).

Now consider the nonequivalence class in the bottom right panel of Figure 20.3. When *tardy* becomes fear-eliciting, so does *late*, which is "the same" as *tardy*. But it would be illogical to expect *early*, the "opposite" of *tardy*, to become fear-eliciting, and research in fact suggests it should not (e.g., Dymond & Barnes, 1995, 1996). In this example, *early* now elicits pleasant responses that are "opposite" to fear. Thus, the function of *tardy* does not "transfer" to *early* but, within the constraints of the relation that unites *tardy* and *early*, it does influence the new function that *early* acquires. The general term that subsumes these kinds of effects is **transformation of function**, which can be defined as the behavioral function of one stimulus changing systematically due to the change in function of another stimulus in the same class (Hayes, Fox, et al., 2001). The transfer of function seen in equivalence classes is a special case of transformation of function, but not all transformation of function is transfer of function.

The critical thing about transformation of function, as illustrated in Figure 20.3, is that, for a given stimulus, alteration of behavioral function will be consistent with the type of relation that unites the stimulus with others. This incorporates many possibilities, including that stimuli *not* united via "same" relations can all acquire the "same" behavioral function, just as happens in equivalence classes. Box 20.1 describes an example. Note that the stimuli in this example are related temporally, and there is nothing inconsistent in responding the "same" way to stimuli regardless of whether they happen "before" or "after."

BOX 20.1

Transformation of Function in a Nonequivalence Class: A Clinical Example

Shawn was a high school student diagnosed with "school phobia." He was rather shy and did not like public attention, but had always enjoyed going to school—until one day a teacher lost her temper and began shouting at him. This frightened Shawn and humiliated him in front of the whole class. In a likely example of classical conditioning, Shawn began to fear his teacher. It was the shouting that had frightened him, but now just glimpsing the teacher upon entering the classroom was enough to cause him to shake and perspire and feel an awful tightness in the pit of his stomach.

Soon, without any additional unpleasant experiences, Shawn also began to fear the classroom itself. Just viewing the other students in their desks from the hallway—even if the teacher was not present—caused Shawn to experience the same unpleasant symptoms. Over time the symptoms were caused as well by seeing the outside of the school building; by seeing the school's sign at the start of a long, curvy drive that led up to the school building; and by hearing the sound of the bus that brought Shawn to school. Before long, Shawn was refusing school entirely, including by leaving the house early in the morning to hide in the neighborhood until after his bus had departed for school. Note that this is not a case of stimulus generalization, involving shared behavioral functions among stimuli that are physically alike (Chapters 17 and 30). The stimuli Shawn feared are as dissimilar as can be. This probably also is not a case of higher-order classical conditioning (in which a neutral stimulus, having been paired with a conditioned stimulus, can elicit the same response as the conditioned stimulus); responding based on this kind of conditioned stimulus tends to be unreliable (Kehoe & Macrae, 1998).

Figure 20.4 summarizes how the stimuli were temporally related in Shawn's experience of typical school mornings. First, Shawn heard the bus approaching his bus stop. The bus picked him up and traveled past the school sign and up to the front of the school building, where Shawn disembarked and proceeded to his classroom. Finally, he entered the classroom and encountered the scary teacher. The various stimuli, therefore, we related in a temporal hierarchy involving what "happens before" and what "happens after."

Shawn's condition likely resulted from the stimuli in the temporal hierarchy being transformed, through direct or indirect association with the scary teacher, to elicit fear. This temporal hierarchy is not an equivalence class. The bus, for instance, isn't the "same as" the teacher, because normally we do not respond in the same way to buses and teachers (we speak to teachers and ride on buses; confusing these behavioral functions would be awkward, to say the least).

How stimuli in a class are functionally transformed depends on the nature of the stimulus relations the class incorporates. In Shawn's case, nothing in the temporal hierarchy

conflicts with a behavioral function of fear. You can be equally afraid of things that "happen before" and "happen after," and so all of the stimuli acquire the same function. But other kinds of relations create other kinds of function transformations. For example, imagine that posted right next to the Sidman High School sign was a second sign indicating the way back to Shawn's home. The two signs would be related as opposites (discussed further in the next section). One is about getting from home to school, and the other is about getting from school to home. Within this relation of opposites, you cannot fear both stimuli. Rather, to remain consistent with the underlying relation of opposites and the learning history that created it, if you fear one sign, then in the presence of the other you should feel the opposite of fear (approach? joy?). The general point is that, within nonequivalence classes, transformation of function can involve either the same or different behavioral functions, depending on the nature of the specific relations that define the class.

Figure 20.4 The temporal-sequence relations among stimuli that became fear-inducing to a school-avoidance student.

SOME TYPES OF NONEQUIVALENCE RELATIONS

We now identify, and provide examples of, a few basic types of nonequivalence relations. All are relevant to everyday experience, but we do not mean to suggest special importance for any of the specific relations that we describe. When harnessing nonequivalence relations for applied interventions, one does not start with a specific type of relation in mind. Rather, the job is to identify specific problems that need solving, and to target the things that people do that require a functional analysis (e.g., Hayes et al., 2001; Rehfeldt & Barnes-Holmes, 2009). The stimulus relations that are implicated will vary from problem to problem, and may even include relations on which no previous research has been conducted. The onus thus falls on the applied behavior analyst to determine what relations are relevant in any specific case and to figure out how to integrate them into an intervention protocol (Rehfeldt & Barnes-Holmes, 2009). The underlying conceptual analysis can be the most laborious part of behavior programming involving derived stimulus relations, but with both equivalence and nonequivalence relations it is critical to precisely define what needs to be directly taught and what untaught relations are logically expected to emerge (e.g., Critchfield & Twyman, 2014). Becoming familiar with several different types of relations will aid in this process.

Distinction Relations

Distinction relations involve responding to stimuli in terms of their differences. In some distinction relations, there is a well-defined basis of stimulus difference. For example, on a dimension of quantity, "many" and "few" are extremes (or at least highly divergent values). Figure 20.5 provides a highly simplified example of a procedure for teaching many/few relations to children (based on Barnes-Holmes, Barnes-Holmes, Smeets, Strand, & Friman, 2004). Here the stimuli are several different types of tokens,[7] and the children learn that these tokens can be exchanged for "many" sweets versus "few" sweets. Relations are taught using a verbal procedure.[8] Thus, to teach AB relations, the teacher says, "This token [points to A] can buy many sweets. This token [A] is opposite to that token [points to B]. Which token would you use if you wanted to buy as many [or few] sweets as possible?" A correct answer is reinforced. To teach BC relations, the teacher says, "This token [points to B] is opposite to that token [points to C]. Which token would you use if you wanted to buy as many [or few] sweets as possible?" In the resulting stimulus class, the function of stimulus C should transform so that it becomes a "many" token that is "the same" as A. Children exposed to this kind of teaching in fact showed the expected derived relations. They were able to mand "more" and "less" using the coins as symbols.

The derived stimulus relations literature tends to divide distinction relations into two subtypes called *opposition* and *comparison* relations, but we are aware of no firm definitions of these subtypes. A reasonable way to think of comparison is as a dimensional distinction relation that does not involve

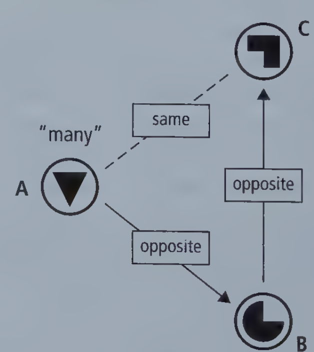


Figure 20.5 A distinction relation.

Based on "Establishing Relational Responding in Accordance with More-than and Less-than as Generalized Operant Behavior in Young Children," by Y. Barnes-Holmes and colleagues, 2004, *International Journal of Psychology and Psychological Therapy*, *4*, pp. 531–538.

"opposites"—but we know of no ironclad rule for deciding when a relation involves "opposites." Sometimes, stimuli that we call "opposite" define endpoints or extremes on some theoretically measurable dimension: "Beautiful" and "ugly" are extremes on a dimension of attractiveness, and "hyperthermia" and "hypothermia" are extremes on a dimension of body temperature. As Box 20.2 illustrates, other "opposites" reflect mutually exclusive instances for which there are no intermediate values.

Recalling that arbitrary relations are defined by the social-verbal reinforcement contingencies that create them, the simplest guide to "opposite" relations may be that often you will know one when you see it. Still, the distinction (pun intended) between opposition and comparison is not always clear cut. For instance, is a many versus few relation (Figure 20.5) really a case of opposites ("extremes on some dimension"), as Barnes-Holmes et al. (2004) described the relations in Figure 20.5? Or is this a comparison relation (after all, "more" is not "most," and "few" is not "least")? We are, frankly, unsure, so we recommend the following perspective to students who are just learning about distinction relations. If good practical reasons exist to teach a particular relation, don't worry too much about what to call it. Instead, expend your energy determining what baseline relations need to be taught to make your desired derived relations a logical outcome (see Luciano, Becerra, & Valverde, 2009).

Spatial and Temporal Relations

Spatial relationships involve the position of a stimulus in reference to the learner or some other stimulus: above versus below, left versus right, in front of versus behind, and so on. **Temporal relationships** involve the relative position of stimuli (or events) in time (e.g., before versus after, per Box 20.1 and Figure 20.4). Spatial and temporal relations are inherent in academic subjects like geography and history, respectively, and these relations are obvious components of everyday nonacademic experience. A child who does not respond appropriately to the parent instruction "Please stand *behind* Tony" or "Eat your ice cream *after* you finish your peas" may soon encounter difficulties. More serious consequences await someone who cannot respond appropriately to the statement "*Behind* you! An axe murderer!"

We are not aware of any applied studies that focus on constructing derived spatial or temporal relations in "pure" form, but several laboratory studies lay out basic procedures that might be translated for applied use (Brassil, Hyland, O'Hora, & Stewart, 2018; May, Stewart, Baez, Freegard, & Dymond, 2017). Spatial and temporal relations also are components in complex deictic relations, which involve relationships between the self and the world, and which have been the focus of considerable research, as we will describe in the "Nonequivalence Relations and Big-Picture Psychological Constructs" section.

Causal Relations

Causal relations are if-then relations. Making accurate causal attributions is an important part of negotiating the everyday world, as illustrated by the prevalence of faulty causal attributions in a variety of clinical problems (Curtis, 1989; Hayes et al., 2001; Matute & Miller, 1998; Torneke, 2010). If-then relations also

BOX 20.2

Opposite Relations in Higher-Order Mathematics

Opposite relations are important to many academic areas of study in the form of either-or categories. Chris Ninness and his colleagues used derived relational procedures to teach college students trigonometry concepts that included cosine and secant, which are reciprocals (Ninness et al., 2009). According to the rules of mathematics, an expression can be cosine or secant, but nothing is intermediate to these two cases.

Figure 20.6 summarizes some of the relations that the students mastered. An interesting feature of this work is that, consistent with what many classrooms require, the learners had to construct a response rather than simply point to a correct stimulus. For instance, shown an example of a cosine (A), they had to write it in reciprocal form (B). This opposite relation was combined with several "same" relations. Given the reciprocal (B), students had to write the equivalent formula for secant (C). Given the secant, they had to construct a graph representing the correct function (D).

Once the baseline relations had been learned, students demonstrated both "same" and "opposite" derived relations. These included the converse of the baseline relations (mutual entailment). For instance, shown the reciprocal of a cosine, they could write the cosine (opposite). Shown a secant, they could write the corresponding cosine reciprocal (same). Also demonstrated were combinatorially entailed relations between stimuli that were not directly associated during training. For example, shown a graphed function, the students could write the relevant cosine formula (opposite) or cosine reciprocal (same). These outcomes generalized to many novel instances, showing that the lesson produced a general skill rather than simply teaching the students about specific examples of cosine and secant.

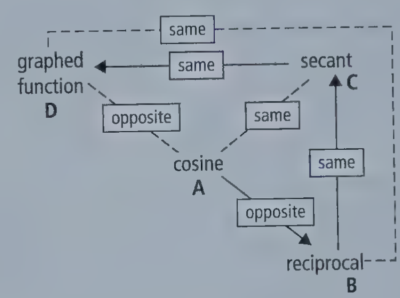


Figure 20.6 A simplified version of same and opposite relations that were established in a college-level trigonometry lesson.

Based on "Constructing and Deriving Reciprocal Trigonometric Relations: A Functional Analytic Approach," by C. Ninness and colleagues, 2009, *Journal of Applied Behavior Analysis*, *42*, pp. 191–208.

are a central feature of understanding and doing science. Just as for other kinds of nonequivalence relations, causal relations can either define the structure of a stimulus class or define the behavioral function through which stimuli in a class are transformed. We are aware of no research on either phenomenon, but in the former case it seems likely that stimuli united by overlapping causal relations (e.g., if A, then B; if B, then C) could cohere into a stimulus class. In the latter case, Box 20.3 illustrates how, for a phenomenon of potential clinical importance, the stimuli in a relational class can be transformed in terms of causal attributions.

BOX 20.3

Learned Helplessness and Transformation of Function

Many clinical conditions are characterized by negative self-evaluations (Curtis, 1989) that a behavior analyst might describe as behavior-controlling rules (Torneke, 2010; see also Chapter 18's treatment of rule-governed behavior). For instance, depressed patients may be heard to remark, "I'm no good at anything" or "Why should I try? I'll just fail" (Peterson & Bossio, 1989). According to learned helplessness theory, these self-evaluative rules are sometimes acquired under conditions of specific unpleasant experiences, during which the individual's behavior is, by definition, ineffective. Then they begin to appear in new situations, where they discourage effective action (Abramson, Seligman, & Teasdale, 1978).

To put it another way, self-evaluative rules can interfere with behavior that could generate important reinforcers (Torneke, 2010). For example, imagine that you are the victim of a mugging. You're walking down the street one night when three people jump out of nowhere, punch you several times, and rob you. The experience is very frightening and, importantly, apparently unpreventable. You were in the wrong place at the wrong time, and nothing you did could have prevented you from being overpowered. Perhaps, as a result of this experience, you say to yourself, "I'm weak and ineffective. I can't even look out for myself." This is a causal rule that can be stated in if-then form: "If placed in a situation where my well-being depends on acting effectively, I will fail." Later, you find yourself being inappropriately passive in many situations. You begin exerting little effort in your classes, which causes your grades to drop. You stop investing in your social relationships, which causes your friends to distance themselves from you. Learned helplessness theory assumes that these things happen because new situations somehow begin to evoke your "helpless" rule (Abramson et al., 1978).

How self-attributions (causal rules) are generated in the first place is a topic unto itself. For now suffice it to say that considerable research shows that humans are prolific in generating them (Barnes-Holmes, Hayes, & Dymond, 2001; Matute & Miller, 1998). Of special interest is how "helpless" self-rules, once formed, become applied to new circumstances. It seems unlikely that this is a matter of stimulus generalization (Chapters 17 and 30), as situations like "being mugged" and "taking a college class" have few physical or perceptual features in common. Thus, something else must be responsible for the spread of inappropriate passivity beyond the situation where unavoidable unpleasant events first were experienced. Cognitive psychologists have proposed that the common factor is a "helpless" mental state (Abramson et al., 1978) that moves with the individual from situation to situation. But Dack, McHugh, and Reed (2009) suggested the common factor as transformation of function involving the original unpleasant experiences and stimuli in new situations.

In the Dack et al. (2009) laboratory investigation, participants first demonstrated the formation of two equivalence classes within a match-to-sample procedure. One class incorporated the word *useless* plus two nonsense words (*lewoly* and *gedeer*), and the other incorporated *good* and two nonsense words (*matser* and *rigund*). Next, participants worked under two different reinforcement contingencies; in only one of these could reinforcement rate be increased by increasing response rate. Using a rating scale, they said that they felt more effective in gaining points on this schedule than on the other one. The two reinforcement schedules were signaled by different-colored lights. For simplicity, let's say that blue signaled the schedule that made people feel effective and yellow signaled the schedule that made them feel ineffective. Finally, in a match-to-sample test, the participants were shown a word (or nonsense word) from the equivalence classes as the sample stimulus, and the two schedule-correlated colors as comparison stimuli. They matched blue to words in the *useless* class and yellow to words in the *good* class.

What do these results show? The fact that participants matched schedule-correlated colors to *useless* and *good* stimuli reflects partly a verbal history (they already knew what those words mean) and partly new learning (they learned the "meaning" of the colors by working on the reinforcement schedules). By assigning the labels *useless* and *good* to schedule-correlated colors, participants basically generated a verbal rule to describe experience with the reinforcement schedules. Through transformation of function, we would expect the nonsense words to also be entangled in rule control. For instance, in a situation where the stimulus *lewoly* is present, the *useless* rule might be occasioned.

This experiment is not a perfect parallel to learned helplessness, but it does show how inappropriate passivity can spread to new situations through pre-existing stimulus classes. By analogy, perhaps your mugging, your school experiences, and your social relations are united in a class that might be called "places where effort and skill matter." Once a "helpless" self-rule arises in the presence of one of these situations, it should arise as well in the others.

Three Practical Considerations

Using Transformation of Function to Create New Reinforcers

Reinforcers are at the heart of just about everything in ABA, but for individuals with highly restricted interests it can be difficult to identify effective reinforcers (Tiger & Kliebert, 2011). In such cases there would be considerable value in a means of reliably creating new reinforcers, and transformation of function can be used for this purpose. Before considering how this works for nonequivalence relations, let us take up the somewhat simpler case of transformation of function in equivalence relations. If an already-effective reinforcer becomes a member of an equivalence class (e.g., using procedures described in Chapter 19), then through transformation of function, other stimuli in the class should also begin to function as reinforcers, even if they have never been directly paired with the original reinforcer. Table 20.1 shows an example (based on Hayes, Kohlenberg, & Hayes, 1991). Imagine that raisins already function as a reinforcer for some individual; audio clips of a particular song do not. The individual is taught to choose a blue triangle when shown a raisin, and also to choose the blue triangle when the audio clips are played. This should establish an equivalence class among the three stimuli, and therefore promote transformation of function. Subsequently, the audio clips (which have never been directly paired with raisins) should function as a reinforcer.[9] New consequences created in this manner have been found effective in changing behavior in standard operant situations such as reinforcement schedules and discrimination learning (Dougher, Hamilton, Fink, & Harrington, 2007; Hayes et al., 1991; Whelan, Barnes-Holmes, & Dymond, 2006).

As the preceding example shows, the behavioral function acquired by new stimuli in an equivalence class is the "same as" that served by an original stimulus. In a nonequivalence class, by contrast, transformation of function is consistent with the types of relations that define the class—sometimes with surprising results. Table 20.1 illustrates (based on Whelan and Barnes-Holmes, 2004). Imagine that pickle juice (P), an already-effective *punisher* for some individual, becomes part of a class incorporating several other stimuli and both "same as" and "opposite" relations (for explanation of how different kinds of relations can incorporate the same stimuli, see the "Theoretical Foundations" section). Two "same as" relations are taught: After smelling pickle juice, choose a blue triangle, and after hearing an audio clip of a song, choose the blue triangle. Two "opposite" relations are taught as well: Shown a picture of a cat, choose the blue triangle as its opposite, and shown a particular doll, choose the blue triangle as its opposite. Through transformation of function, the audio clip becomes a punisher (it was "same as" the blue triangle, which was "same as" the punisher). Of special importance to the present discussion, the cat and the doll became *reinforcers* (they were opposites of the blue triangle, which was "same as" the punisher). Thus, new reinforcers are created even though no other stimulus in the nonequivalence class initially functioned as a reinforcer.

Do not feel bad if you find the examples of Table 20.1 to be a little mind-bending, because standard thinking about reinforcers has not prepared you for them. The usual story told about conditioned reinforcers is that they result from the pairing of neutral stimuli with already-effective reinforcers (see Chapter 11). But the stimuli that become learned reinforcers in Table 20.1 were never directly paired with an already-effective reinforcer, suggesting that a pairing-focused account of conditioned consequences is too narrow. This is not surprising given that the pairing view arose long before the start of systematic research on derived stimulus relations (e.g., Hull, 1943; Skinner, 1938; Wyckoff, 1959). But even more recent theoretical accounts of conditioned consequences are silent on derived effects (e.g., Fantino, Preston, & Dunn, 1993; Grace & Savastano, 2000; Shahan, 2010). Thus, it is not exaggerating to say that, in this area at least, research on derived stimulus relations forces a reconsideration of the very fundamentals of behavior theory.

Returning to the mission of the present book, the effects described in Table 20.1 suggest a way to use derived stimulus

TABLE 20.1 Creating New Reinforcers Through Transformation of Function Within Equivalence and Nonequivalence Classes

Stimulus	Paired With	Type of Relation	Original Function	Transformed Function
Equivalence Class				
Raisin	Blue triangle	Same	Reinforcer	—
Audio clip	Blue triangle	Same	—	Reinforcer
Nonequivalence Class				
Pickle juice	Blue triangle	Same	Punisher	—
Audio clip	Blue triangle	Same	—	Punisher
Cat	Blue triangle	Opposite	—	Reinforcer
Doll	Blue triangle	Opposite	—	Reinforcer

Equivalence class example: Based on "The Transfer of Simple and General Consequential Functions Through Simple and Conditional Equivalence Relations," by S. C. Hayes and colleagues, 1991, *Journal of the Experimental Analysis of Behavior, 56*, 119–137. Nonequivalence class example: Based on "The Transformation of Consequential Functions in Accordance with the Relational Frames of Same and Opposite," by R. Whelan and colleagues, 2004, *Journal of the Experimental Analysis of Behavior, 82*, 177–195.

relations to create potentially endless reinforcers for any individual. Much remains to be learned, however. For example, to our knowledge the long-term durability of learned reinforcers established as in Table 20.1 has not been explored (we assume that the new reinforcers would be subject to the same motivating operations as the original reinforcer, and that derived reinforcing effects would depend on retention of the stimulus class on which transformation of function is based). This is an interesting direction for future research and application.

Incoherent Stimulus Classes

So far we have described derived stimulus relations in which the directly learned baseline relations predict the exact nature of derived relations. However, this will not be the case for every set of overlapping conditional discriminations. Consider an example involving distinction relations, summarized in Figure 20.7 (left panel), in which it is unclear on what precise basis the stimuli differ:

> If you are told that "bees are different from mammals," you cannot know what bees are like, or exactly which features differentiate them from mammals. Furthermore, the derived combinatorially entailed difference relations also are unspecified. For example, if you are told that bees are different from mammals, and that mammals are different from birds, you cannot know how bees and birds differ, or whether they are even the same (Luciano et al., 2009, p. 169).

Figure 20.7 (right panel) shows another example. A child who is just starting to learn about money is taught that *25¢* is "more than" *10¢* and also that *10¢* is "more than" *5¢*. These relations, by themselves, do not provide any basis for judging whether *10¢* or *25¢* is "more" (note that the issue is not what is true, but what can emerge from the already-mastered relations).

Also potentially undefined in classes like these is how transformation of function might proceed. For instance, in the left panel of Figure 20.7, what happens if you are attacked by a pugnacious parrot and thereafter fear birds? How (if at all) does this affect your emotional response to bees and mammals? *Something* may happen (stimulus control will always exist, and when experimental control is weak, unknown influences from the subject's learning history will predominate; Sidman, 1960), but in this case, based on the directly taught relations in the stimulus class, we cannot guess what.[10]

These examples highlight why we have stressed the importance, in derived stimulus relations interventions, of carefully mapping out the relations to be taught and the relations that logically should emerge from this teaching. A defining advantage of applied work involving derived stimulus relations is the "free learning" (in the form of untaught relations) that arises from it (e.g., Critchfield, 2018; Rehfeldt, 2011; Rehfeldt & Barnes-Holmes, 2009). From this perspective, applied relational interventions are pointless if one cannot specify the derived relations that training is expected to yield.

Tools for Planning Interventions Involving Nonequivalence Relations

Historically, an applied behavior analyst interested in teaching derived stimulus relations had to create stimuli and devise training protocols from scratch. Recently, however, efforts have been made to develop standardized tools to ease this process. We will mention two.

Training and Assessment of Relational Precursors and Abilities (TARPA) is a free software program for assessing and establishing derived stimulus relations. Several studies support its reliability and validity (Moran, Stewart, McElwee, & Ming, 2010, 2014; Moran, Walsh, Stewart, McElwee, & Ming, 2015). As of this writing, the program was available for equivalence relations only, but expansion into nonequivalence relations was under development. For information, go to http://www.vb3.co.uk/

The *PEAK* (Promoting the Emergence of Advanced Knowledge) *Relational Training System* is a commercially available assessment and treatment curriculum for adults and children with autism and related disabilities. Each volume contains 184 structured lessons and related assessment tools. One of the four volumes focuses on derived equivalence and another focuses on derived nonequivalence relations. More than 20 peer-reviewed studies provide evidence of the PEAK curriculum's reliability, validity, and effectiveness (for a review, see Dixon et al., 2017).

THEORETICAL FOUNDATIONS

The examples of the last section were intended to clarify the "what" of derived nonequivalence relations. In the present section we take up the "why" of those relations, in two ways. First, in a discussion of *contextual stimuli*, we explain how relational behavior, like all operant behavior, is situation specific. Second, we introduce a leading theory of derived relations, with emphasis on two of its most interesting propositions.

Stimulus Relations Are Defined by Contextual Control

To repeat an observation from earlier, in arbitrary stimulus relations nothing about the stimuli themselves dictates how they can be related. Rather, human social/verbal customs and reinforcement contingencies decide this. These customs and contingencies exert their influence through **contextual stimuli**, which

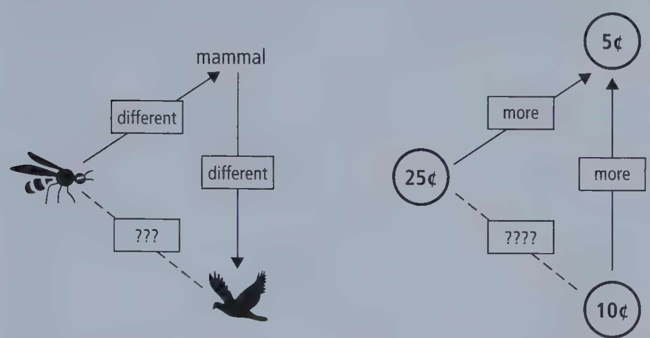


Figure 20.7 Two logically incoherent stimulus classes. Both include overlapping conditional discriminations, but the directly learned baseline relations do not support firm predictions about combinatorial entailment.

may be said to signal the specific way in which stimuli are to be related (Bush, Sidman, & Rose, 1989; Hayes, Fox et al., 2001). Figure 20.8 shows how the relations among a single set of stimuli might change depending on contextual cues. In this example (adapted from Berens & Hayes, 2007), a child must master the relations among unfamiliar coins, including a U.S. quarter (*25¢*), dime (*10¢*), and nickel (*5¢*). In the left panel, when the child is shown *25¢* and then asked, "Which is less?" a correct response is to identify *10¢*. In similar fashion the child can be taught that *5¢* is "less than" *10¢*. This should yield a derived relation in which the child, shown *10¢* and allowed to select either *5¢* or *25¢*, correctly identifies *5¢* as "less." Everything reverses, however, when the context is "more than" (center panel). And given the context of "money," the coins could be treated as equivalent (right panel; they are all types of money).[11] Thus, the same stimuli can be related in any number of ways, depending on the contextual cues.

In a sense, contextual stimuli signal what a person "needs to know" (that is, what way of relating stimuli will be reinforced) in a given situation. To get a better sense of contextual stimuli, imagine that you are completing a match-to-sample task, and the sample stimulus is *Picasso*. The comparison stimuli are *Spain*, *Michelangelo*, and *Gris*. Which one should you choose? Without knowing "the question being asked," this is an impossible decision, because each of these comparisons is an appropriate match for *Picasso* under some circumstance. *Picasso* was born in *Spain*. Like *Michelangelo*, he did some sculpting. Like Juan *Gris*, he used the painting style known as analytical cubism. Here "the question being asked" is a contextual cue that tells you on what basis you should choose a comparison stimulus. Because learned contextual cues mediate the nature of arbitrary stimulus relations in this way, contextual stimuli are an important component of applied stimulus relations interventions.

In many of the examples provided in this chapter, we mention verbal contextual stimuli (e.g., "sculpting," above) because these are easy to understand. But contextual stimuli don't have to be verbal. Imagine that you live in a lovely house, but your hobby involves survivalist adventures in which you must live off the land in the wilderness. In the context of being at home, insects fall into a class of stimuli that are to be avoided—for instance, you might call an exterminator to rid your house of termites. In the context of being alone in the forest without food, however, termites, as a source of life-saving nourishment, may fall into a class of things to be sought rather than avoided. A contextual stimulus is merely one whose presence correlates with contingencies favoring a particular way of relating.

Are Contextual Stimuli Always Essential?

The preceding sets the stage for a critical look at match-to-sample procedures for establishing derived stimulus relations like those described in Chapter 19. In much of the published research on stimulus equivalence that has been conducted to date, you will be hard pressed to find reference to contextual stimuli as we are discussing them here.[12] The example described in Figure 19.1 illustrates. During the training of each baseline relation, the sample stimulus was a spoken word, and the comparison stimuli were pictures. Thus, on one type of trial, the sample might be "car" and the comparisons might be pictures of a dog, a car, a cow, and a coffee mug. Selecting the car was reinforced. At no point during this procedure, however, was there any discrete stimulus whose presence correlated with the "same" contingency. Earlier we said that derived stimulus relations depend on contextual stimuli to signal the type of relation that should operate. Many different relation types could unite the stimuli. For instance, a coffee mug is something you might take with you in a car, "cow" starts with the same sound as "car," and a dog is something that chases a car.

How, then, did the participants know what to choose from trial to trial? Two possibilities exist. The first is that, in the relatively simple confines of the experiment (and unlike in the everyday world) only one type of relation was ever reinforced, so there was no need for contextual cues. In the experiment of Figure 19.1, once "same" responding was shaped up, it applied equally well (i.e., was reinforced) across all parts of the experiment.

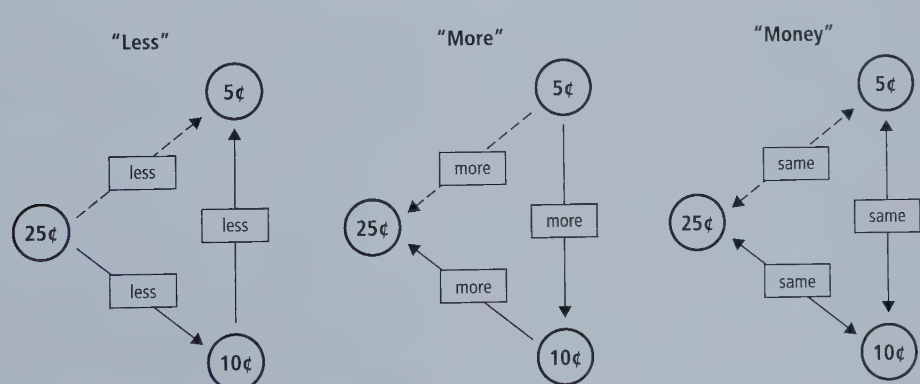


Figure 20.8 How contextual cues define stimulus relations. Left: In a "less" context, coins are related in a certain way (arrows point from the more valuable to the less valuable coin). Center: In a "more" context, those relations are reversed. Right: In the context of "money," the coins are essentially "the same."

Left and center panels are based on "Arbitrarily Applicable Comparative Relations: Experimental Evidence for a Relational Operant," by N. M. Berens and S. C. Hayes, 2007, *Journal of Applied Behavior Analysis*, *40*, 45–71.

A second possibility is that the experiment actually contained contextual stimuli, but they were implicit, meaning the overall experimental setting became a compound contextual stimulus that was correlated with reinforcement contingencies. In effect, perhaps the participant learned, "When I'm here, in this room, with this experimenter and these stimuli, I should respond with 'same.'" There is plenty of precedent for contextual cues of this type. In the classical conditioning literature, for instance, it is well established that contextual stimuli can be compounds of discrete stimuli plus aspects of the setting where conditioning took place (Bouton & Nelson, 1998).

Neither of these possibilities bodes particularly well for the generality of what is learned (Chapter 30). Someone whose relational responding is not under contextual control will be at a loss for what to do in a world where the contingencies are always changing (it will not pay, for example, to point to a picture of a car when asked, "What do you take with you in a car?"). Someone whose contextual stimuli for "same" responding are highly specific to one training situation may not make "same" responses when somewhere else.

What all this suggests to the applied behavior analyst is that efforts to create derived stimulus relations should, whenever possible, incorporate explicit, naturalistic contextual stimuli in the form of cues that are likely to be present in everyday environments. In the human world, these cues will often be verbal in nature—hence the many verbal examples employed in the present chapter. While it may be possible to build an equivalence class using a strictly nonverbal match-to-sample procedure, for verbal learners we recommend including a contextual prompt such as "pick the one that is the same," or simply, "same."

Relational Frame Theory

Now that you have a sense of what nonequivalence relations are, it will be helpful to briefly examine the leading theoretical account of them, **relational frame theory** (RFT; Hayes et al., 2001), which you encountered briefly in Chapters 18 and 19. In the present chapter we will not take up the question of whether RFT or some other theory of stimulus relations, like Sidman's (2000), is "better." However, RFT has had sufficient impact on research into derived stimulus relations that any student who wishes to become familiar with this topic needs a working understanding of the theory. To date, the vast majority of studies conducted on nonequivalence relations have been inspired by RFT, and as a result the vocabulary of nonequivalence relations used in the present chapter (e.g., *mutual entailment* and *combinatorial entailment*) comes from RFT. Additionally, although there is wide agreement that derived stimulus relations provide a basis for understanding "higher cognitive abilities" of humans (e.g., Hayes et al., 2001; Sidman, 1986, 1994), RFT researchers have been especially focused on this possibility (e.g., see the "Nonequivalence Relations and Big-Picture Psychological Constructs" section). A complete explanation of RFT is beyond the scope of the present discussion, but we will mention two key features of the theory that will help in understanding points to be made later in the chapter.

Proposition 1: The Behavior of Relating Stimuli Is Verbal Behavior

RFT assumes that derived stimulus relations are inherently verbal relations (Hayes, 1984).[13] That is, the behavior of relating stimuli *is* verbal behavior. We know of no evidence that directly confirms or refutes this assumption—it may not even be definitively testable—but it is supported by circumstantial evidence. For example, derived stimulus relations have, at best, been produced in nonhumans only rarely and as the product of considerable effort (e.g., Galizio & Bruce, 2018). This stands in stark contrast to the apparent ease with which verbally proficient humans derive stimulus relations and the broad flexibility with which they apply them (e.g., Figure 20.1). Moreover, as we will describe later in the chapter, human relational abilities appear to arise during the same period of child development in which verbal behavior grows precipitously. These observations suggest a profound linkage of derived stimulus relations and the unique verbal capabilities of humans.

Thinking of stimulus relations in terms of verbal behavior raises the interesting question of what is going on in match-to-sample protocols for establishing stimulus relations, which Chapter 19 portrayed as largely nonverbal procedures. One possibility is that in "nonverbal" circumstances learners apply covert verbal behavior to master stimulus relations (e.g., Horne & Lowe, 1996; see Chapter 19 for information about naming theory). For present purposes, however, we merely consider the possibility that, if the relating of stimuli is a verbal process, then verbal behavior can be strategically harnessed to construct stimulus relations. A number of studies show that baseline relations can be established, and derived relations promoted, by simply telling the learner about the required overlapping conditional discriminations (e.g., Berens & Hayes, 2007; Smyth, Barnes-Holmes, & Barnes-Holmes, 2008). This suggests a way, in applied work, to established stimulus relations more efficiently than through laborious match-to-sample protocols, as Box 20.4 discusses. Note that research of this type foreshadows other ways in which verbal interactions may intersect with derived stimulus relations, a topic that we will take up in this chapter's final section.

Proposition 2: Reinforced Relational Experience Yields a Generalized Relational Repertoire

Chapter 19 described equivalence classes as the result of reinforced practice with baseline relations involving specific stimuli. For example, consider the class in Figure 20.9 (top panel). According to the standard view, acquiring this class requires two kinds of baseline learning experiences: reinforced experience selecting $p \leq .05$ (rather than $p > .05$) in the presence of *low p value*, and reinforced experience selecting *statistically significant* in the presence of *low p value* (selecting the opposite answers in the presence of *high p value* also would need to be reinforced). Also according to the standard view, acquiring each new equivalence class (e.g., *overdue*, *tardy*, and *late*; top right panel of Figure 20.3) would require the analogous steps. RFT, however, proposes a developmental process that alters the nature of stimulus-class acquisition. At the core of this proposal is a disagreement with Sidman's (2000) theory

BOX 20.4

Telling Isn't Teaching—Right?

About that much-maligned mainstay of classroom teaching, the lecture, it has long been observed that "Telling isn't teaching." The corollary is that passive listening isn't learning, as Skinner (1958) suggested when advocating for teaching machines: "Unlike lectures . . . the machine induces sustained activity. The student is always alert and busy" (p. 971). In behavioral systems of instruction (e.g., Keller, 1968; Skinner, 1958), student learning is said to depend on frequent, active responding with feedback. This has been suggested specifically about equivalence-based instruction (e.g., Critchfield & Twyman, 2014).

It would therefore seem unlikely that derived stimulus relations could be established simply by telling learners about the baseline relations. But consider the case of some college students who learned about inferential statistics through an automated lesson in an online course delivery system (Critchfield, 2014). At each step of the lesson, the students were asked to read "notecards" presenting brief statements such as:

> With a low *p* value, the results are statistically significant. With a high *p* value, the results are not statistically significant.

The preceding statements incorporate two AB relations (one of which is diagrammed in Figure 20.9, top panel). After reading each statement, a student answered two to four multiple-choice questions such as

Low *p* value

A. not statistically significant
B. statistically significant

These questions were intended to serve as an attention check. Correct answers earned the feedback "Correct," and incorrect answers earned the feedback "Wrong," along with a prompt to review the relevant notecard. After this exposure to the baseline relations AB and BC, between 89% and 100% of the students answered test questions correctly about expected derived relations (BC; right panel of Figure 20.9). For instance, relevant to the relations shown in the left panel, students correctly answered this question:

$p \leq .05$

A. statistically significant
B. not statistically significant

The noteworthy feature of this lesson was its efficiency. With "telling" substituted for the elaborate sequences of match-to-sample trials that are found in many stimulus relations protocols, students completed the lesson in about 25% to 50% of the time previous students needed to master the same concepts in a match-to-sample protocol (see Fienup & Critchfield, 2010, 2011). In this particular instance, perhaps "telling" *was* teaching. Perhaps. There is uncertainty on this point because the attention-check questions did require some active responding, and even though students did not have to answer these

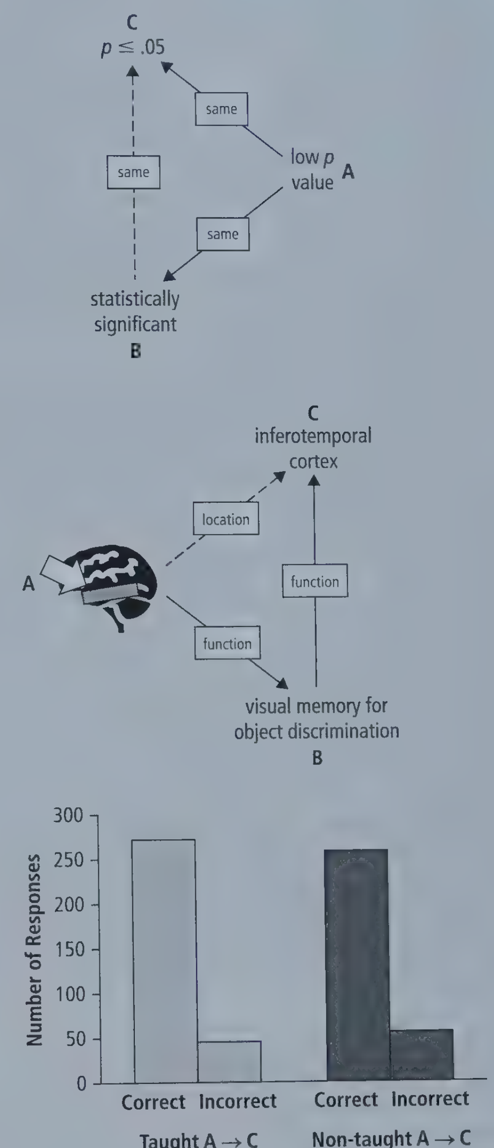


Figure 20.9 Stimulus relations established in college students by telling about baseline relations rather than reinforced match-to-sample experience. Top: Concepts related to inferential statistics. Middle: Concepts related to neuroanatomy. Bottom: Correct responses on one relation for students directly taught it versus those expected to demonstrate it emergently. For each group, separate bars show correct responding for students who did and did not first show mastery of baseline relations.

Top based on "Online Equivalence-based Instruction About Statistical Inference Using Written Explanation Instead of Match-to-Sample Training," by T. S. Critchfield, 2014, *Journal of Applied Behavior Analysis*, *47*, 606–611; middle based on "Using Stimulus Equivalence Technology to Increase Efficiency in Teaching Neuroanatomy," by C. L. Pytte and D. M. Fienup, 2012, *Journal of Undergraduate Neuroscience Education*, *10*(2), A125–A131; bottom "Using Stimulus Equivalence Technology to Increase Efficiency in Teaching Neuroanatomy," by C. L. Pytte and D. M. Fienup, 2012, *Journal of Undergraduate Neuroscience Education*, 10(2), A125–A131, Reproduced by permission from the Pytte and Fienup study.

questions correctly (there was no mastery criterion), responses did produce contingent feedback. Thus one cannot rule out that active responding and feedback, even in small doses, were necessary to the study's good outcomes.

Here a study by Pytte and Fienup (2014) is potentially informative. College students learned about neuroanatomy concepts involving brain areas, their locations in the brain, and their behavioral functions. The medium of instruction was a lecture organized to emphasize certain baseline relations. The center panel of Figure 20.9 shows one example. For seven brain regions, the AB and BC relations were explained in the lecture, but the expected derived AC relations were not. For seven other brain regions, which served as positive controls, the AC relation also was explained.

On the course exam, students averaged approximately 82% correct on questions about untaught AC relations, regardless of whether or not those relations were addressed in the lecture. Thus, "telling" about baseline relations was sufficient to promote combinatorial entailment. This is not really surprising, because past laboratory research had shown similar effects from "telling" about baseline relations (Smyth et al., 2008), and from simply "demonstrating" them by presenting pairs of stimuli without any requirement for overt responding (e.g., Leader & Barnes-Holmes, 2001a).

What is the "best" approach to establishing stimulus classes? Although "telling" and "demonstrating" alone apparently can work under at least some circumstances, perhaps even better outcomes would be obtained by requiring active responding with feedback, per the Critchfield (2014) study. On this point, however, the evidence is mixed. In comparisons of the "demonstration" procedure with match-to-sample, stimulus classes were more reliably established via the latter in one study (Clayton & Hayes, 2004) and by the former in another (Leader & Barnes-Holmes, 2001b). In at least one study with college students learning about single-subject designs, "telling" via lecture and a match-to-sample protocol promoted derived stimulus classes equally well (Lovett, Rehfeldt, Garcia, & Dunning, 2011).

To help make sense of these findings, it is worth noting an especially interesting feature of the Pytte and Fienup (2014) study that is summarized in the bottom panel of Figure 20.9. For both taught and untaught AC relations, Pytte and Fienup calculated accuracy separately for students who did, and did not, correctly answer questions about baseline relations. Regardless of whether students were told about AC, the critical issue about AC relations seemed to be that baseline AB and BC relations first were thoroughly mastered. For students who *did* acquire the baseline relations, telling was not necessary for combinatorial entailment. For students who *did not* master baseline relations, telling was insufficient for combinatorial entailment.

The Pytte and Fienup (2014) results suggest that, as with match-to-sample protocols, the critical thing about "telling" or "demonstrating" procedures may be whether they firmly establish baseline relations. Unfortunately, although "telling" and "demonstrating" have been staples of classroom instruction for centuries, the conditions under which these strategies reliably change behavior remain to be fully understood. It is therefore prudent, for the time being, to follow the lead of most work on behavioral systems of instruction (e.g., Keller, 1968; Skinner, 1958) and assume that active responding and contingent feedback, in some measure, are useful features of stimulus relations protocols. This does not, however, negate the potential of other procedures to greatly streamline the acquisition of stimulus relations by learners with adequate verbal skills.

asserting that derived relations are an automatic outcome of specific reinforcement contingencies. RFT assumes instead that, given the right learning history, the deriving of stimulus classes becomes a generalized repertoire that RFT calls *arbitrarily applicable relational responding* (Barnes-Holmes & Barnes-Holmes, 2000) that potentially applies to any number of specific stimuli.

The required learning history is reminiscent of that required to create generalized imitation (Baer, Peterson, & Sherman, 1967). For at least some learners, a teacher's command to "Do this" initially is insufficient to prompt the copying of an action like touching one's nose. Direct teaching can bring that particular action under the control of "Do this," but the prompt initially may not work with other actions (e.g., clapping). Given direct instruction of imitation of enough actions, "Do this" becomes a generalized contextual cue that prompts the copying of any action under the stimulus control of what the teacher does next. In this example you may recognize **multiple-exemplar training**, which was described in Chapter 17 as the repeated acquisition of some behavioral repertoire with variations in situations, stimuli, or response features (see also Holth, 2017). In the RFT account of generalized relational operants, each new stimulus class of a particular type (like equivalence) that reinforcement teaches is one exemplar, and given enough reinforced class formations, the deriving of stimulus classes becomes a higher-order operant that potentially applies to any number of specific stimuli. It does this through the contextual stimuli that were mentioned earlier in the chapter. In short, what multiple-exemplar training creates is generalized control by contextual stimuli.

When contextual stimuli become generalized, the relational behavior they mediate is called **arbitrarily applicable relational responding** (AARR; Hayes, Fox et al., 2001), which means forming new stimulus classes with little to no new reinforced practice. AARR involving any specific type of relation is called a **relational frame** (Hayes, Fox et al., 2001) and is thought to be acquired separately from that involving other types. A possible example of AARR is the following: When you see a supermarket sign indicating that the store's price on your favorite snack is "less than" the price at another store, you don't need any new learning experiences to know how to respond to that.

Because AARR is thought to result from accumulated multiple-exemplar experience, very young children, who lack an extensive history of reinforced relating, would be expected *not* to spontaneously demonstrate derived relations after learning baseline relations. Rather, they would need to acquire those relations by learning about baseline relations through reinforced practice. As relational experience accumulates, however, children should begin to form stimulus classes without extensive reinforced conditional discrimination training.

This perspective has two important implications. The first concerns developmental trajectories: Normatively speaking, in childhood, relational abilities should improve with age, although not necessarily at the same pace for all types of relations because each type is its own generalized class requiring its own formative experiences. In the next section, we describe evidence that supports this prediction. The second implication is that, using multiple-exemplar training, it should be possible to experimentally create a particular type of AARR. Numerous studies have shown that when enough exemplars of one type of stimulus relation are directly trained, new relations of the same type may subsequently be formed without further training (Barnes-Holmes et al., 2004; Berens & Hayes, 2007; Dunne, Foody, Barnes-Holmes, Barnes-Holmes, & Murphy, 2014; Heagle & Rehfeldt, 2006; May et al., 2017; Ninness et al., 2009).[14] Collectively, these studies remind us that a derived stimulus relations intervention can have two goals: to establish a stimulus class involving specific stimuli, and to establish a generalized repertoire through which the learner might form similar classes with new stimuli.

A brief editorial comment is in order: You needn't be a fan of RFT to appreciate studies showing generalized repertoires of relating. One of the most difficult challenges in ABA is to create behavior change that generalizes beyond the confines of a particular intervention setting (Stokes & Baer, 1968; Chapter 30). AARR may well be this kind of effect, although much remains to be learned about it. To date, research has shown mostly that, given multiple-exemplar experience, newly acquired relational repertoires can apply to novel stimuli in the same setting where an intervention took place. It is quite another matter to suggest that relational abilities acquired in one setting extend to very different ones (e.g., clinic-established relating also appears at home or at school). We are not aware of any research conducted to date that applies this stricter test of the AARR concept.[15] Nevertheless, findings like those just described underscore that it may be possible with stimulus relations interventions not just to establish a particular stimulus class of interest but also to set generalized repertoires of relating into motion.

NONEQUIVALENCE RELATIONS AND BIG-PICTURE PSYCHOLOGICAL CONSTRUCTS

As many of the diagrams in this chapter illustrate, behavior analysts often think about behavior in terms of discrete stimuli and responses, often as these unfold moment by moment. Outside of behavior analysis, people often operate at a different level of analysis. They speak of big-picture constructs like "intelligence" and "knowledge" that refer to functioning over extended time frames. Skinner (1945) saw a means of reconciling the two levels of analysis by viewing psychological constructs as umbrella terms for potentially complex behavioral phenomena. If you want to know what "intelligence" is, Skinner suggested, look at the various behavioral circumstances under which people speak of "intelligence." In his view, psychological constructs tact *something*. That something is behavior, and it can be informative to specify what kind of behavior.

From the outset, pioneering researchers of relational learning saw their mission as providing an alternative to mainstream thinking about big-picture psychological constructs (Hayes, 1984; Hayes et al., 2001; Sidman, 1986, 1994). In the present section, we briefly explore some possible connections between derived stimulus relations and three such constructs: knowledge structures, self-concept, and intellectual development.

Categorical Knowledge: Hierarchical Relations

We opened the present chapter by referring to semantic networks as hypothesized structures that organize bits of knowledge in long-term memory. Figure 20.1 showed that many different kinds of derived stimulus relations may be implicated in the behavioral phenomena that lead observers to speak of "knowledge" in the first place. One such case involves responding to stimuli in categories, which are classes within which (a) nonidentical stimuli are treated as "same," and (b) the same stimuli all are simultaneously "different" from stimuli in other classes. As Lakoff (1987) has written:

> There is nothing more basic than categorization to our thought, perception, action, and speech. Every time we see something as a *kind* of thing . . . we are categorizing. Whenever we reason about *kinds* of things—chairs, nations, emotions, any kind of thing at all—we are employing categories. . . . An understanding of how we categorize is central to any understanding of how we think and how we function, and therefore central to an understanding of what makes us human. (pp. 5–6)

Category relationships, by their nature, involve nonequivalence, and thus hierarchical relations have much in common with distinction relations. Though related to *ale*, *beer* is not the "same as" *ale*. Rather, *beer* is a category that includes *ale* (Figure 20.1), but *ale* does not include *beer*. Conversely, *ale* is a member of the category *beer*, but *beer* is not a member of the category *ale*. Such asymmetrical stimulus relations in categorization have been described as **hierarchical relations**, which involve a stimulus "including" (serving as a category label for) other stimuli, and those stimuli being "member of" the category stimulus. Hierarchical relations imply a nesting of multiple classes of stimulus relations. Thus, *India pale ale* and *English bitter*, though "different" in some ways, are "the same" in that they are members of the stimulus class *ale*. *Ale* and *lager*, though "different," are "the same" in being types of *beer*. *Beer* and *wine*, though "different," are "the same" in being types of *alcoholic beverages*, which are "different" from *nonalcoholic beverages*, though both the *alcoholic* and *nonalcoholic* varieties are types of *beverages*—and so on.

We will not attempt to explain the procedures by which derived stimulus relations researchers have built hierarchical relations because these can be quite complex (e.g., Griffee & Dougher, 2002; Paliliunas, Belisle, Barron, & Dixon, in press). A good place for the interested reader to begin may be with general considerations that apply to establishing hierarchical relations (Luciano, et al., 2009), one of the most important of which is establishing contextual stimuli that mediate whether two stimuli are related via "member of" or "includes" relations (e.g., Slattery & Stewart, 2014). For instance, in the context of "includes," *animal* "includes" *dog* (which is a type of animal), but *dog* does not "include" *animal* (which is not a type of dog). In the context of "member of," *dog* is a "member of" *animal*, but not vice versa. Once established, the contextual stimuli can assume generalized properties such that they guide the categorization of novel stimuli (Gil-Luciano, Ruiz, Valdivia-Salas, & Suárez-Falcón, 2017; Slattery & Stewart, 2014).

To summarize thus far, the research literature shows that derived stimulus relations protocols can be used to establish relations with hierarchical properties, but this does not necessarily show that similar processes operate when people learn to categorize in the everyday world. Real-world categorization typically exhibits three effects (Murphy, 2002) that any approach based on derived stimulus relations must reproduce to be credible. Box 20.5 introduces the phenomena and illustrates how hierarchical categories can be conceptualized in terms of derived

BOX 20.5

Three Characteristics of Hierarchical Categories

Researchers outside of behavior analysis have found that three properties reliably apply to the way that people categorize (Murphy, 2002). Two of these 'containment,' referring to what stimuli are subsumed by which categories, and the third concerns "property induction," which may be thought of as the transformation of behavioral functions.

Asymmetrical class containment refers to the fact that a higher-order class (e.g., "*animal*") contains a lower-order class (e.g., "dog") but not vice versa (i.e., "dog" does not contain "animal") (Slattery & Stewart, 2014, p. 61).

Figure 20.10 (top panel) shows that contextual stimuli controlling "includes" and "member of" relations can be thought of as dictating unidirectional stimulus control. The figure is arrayed such that in the "includes" context an item includes only items positioned to its right. In the "member of" context, an item is only a member of items positioned to its left. This applies equally to relations that are directly taught (e.g., in the figure, that *animal* "includes" *dog* and *ant*) and derived (e.g., *animal* includes *poodle*) relations.

Transitive class containment refers to classifying a stimulus (A) as a member of a higher-order class (C) on the basis that it is a member of a subclass (B) that is a member of that higher-order class. For example, if a child is taught that "poodle" (A) is a type of "dog" (B), then the child may also classify "poodle" as an "animal" (C) on the basis that "dog" (B) is a type of animal (Slattery & Stewart, 2014, p. 61).

The derived stimulus control of transitive class containment flows unidirectionally in accordance with contextual cues. Because *animal* "includes" *dog* and *dog* "includes" *poodle*, in Figure 20.10 *animal* is understood to "include" *poodle* even if this relation is not directly taught. In similar fashion, *poodle* would be understood as "member of" *animal*.

Unilateral property induction refers to the concept that properties or features of a higher-order class (e.g., "animal") will also be found in a lower-order class (e.g., "dog") but not vice versa. For example, all animals breathe and thus dogs breathe; however, while dogs have four legs, not all animals do (Slattery & Stewart, 2014, p. 61).

Figure 20.10 (bottom panel) depicts unilateral property induction as transformation of function flowing unidirectionally in accordance with the contextual cues. An individual who learns that "breathes" is a property of *animals* will understand that *dogs* breathe and *poodles* breathe. An individual who learns to identify "curly hair" as a property of *poodles* will not assume that all *dogs* or all *animals* share this property.

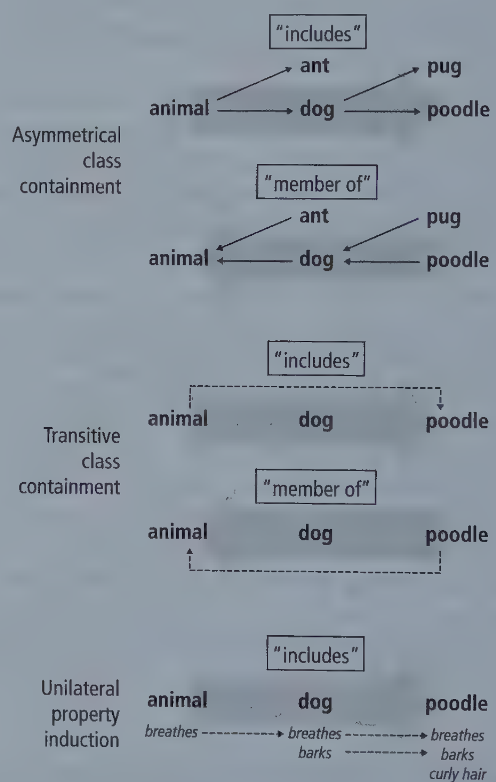


Figure 20.10 Examples of derived stimulus relations inherent in three effects commonly seen in hierarchical categorization. Note that, unlike in other diagrams in this chapter, arrows do not point from sample to comparison stimulus. Rather, arrows indicate class containment; for instance, for the contextual cue "includes," an arrow pointing from *animal* to *dog* indicates that the category *animal* includes *dog*.

stimulus relations. Recent studies have shown that these phenomena do in fact arise when hierarchical relations are created via derived stimulus relations protocols (Slattery & Stewart, 2014; Slattery, Stewart, & O'Hora, 2011).

Self-Concept: Deictic Relations

Much of this chapter has described how people relate different aspects of the world to each other. But people also focus a great deal on the relations between themselves and the world. Various psychological constructs such as identity, self-concept, and self-image at least partly address this self/world relation. As is true for *beer* (Figure 20.1), however, a "self" cannot be understood outside the context of things to which it is related. RFT provides a detailed theoretical account of various aspects of this "self" and the relations through which it arises and manifests (Barnes-Holmes, Hayes, & Dymond, 2001; Torneke, 2010). As we will note briefly in the chapter's final section, this account is useful in understanding and resolving a variety of clinical problems.

Let's start to unpack "sense of self" by considering how "self" is related to other people. Asked what her dog in the next room sees, a young child may instead describe what *she* sees. The child does not understand that one's vantage point is different from that of others. **Perspective shifting**[16] means responding as if you see the world as from another's vantage point, or perceive a situation as if were you in another's position (Batson, Early, & Salvarani, 1997). But there are other types of perspective shifting. Very young children may also confuse the perspectives of now with then (asked, "What did you do yesterday?" the child may describe what she is doing now) or the perspectives of here and there (asked, "What did you see at Grandma's?" the child may describe what she sees in her own house).

Perspective-shifting abilities correlate with general social success (e.g., McHugh, Barnes-Holmes, Barnes-Holmes, Stewart, & Dymond, 2007), and disruptions in this ability are associated with reduced "emotional intelligence" (Bar-On & Parker, 2000) and several clinical conditions (e.g., Ingram, 1990; Mullins-Nelson, Salekin, & Leistico, 2006), including autism (Baron-Cohen, Leslie, & Frith, 1985). Additionally, as this chapter's final section explains, perspective-shifting repertoires are relevant to certain problems of adjustment experienced by typically developing individuals.

In the behavior analytic view, perspective shifting involves specific kinds of stimulus relations. These are learned relations rather than the automatic product of maturation (Weil, Hayes, & Capurro, 2011), and the everyday environment does not always provide the needed learning experiences (McHugh, Barnes-Holmes, & Barnes-Holmes, 2009; Valdivia-Salas, Luciano, Gutierrez-Martinez, & Visdomine, 2009). Fortunately, research is beginning to show how to promote the acquisition of selected stimulus relations related to perspective shifting, for both typically developing children (Heagle & Rehfeldt, 2006; Montoya-Rodriguez & Cobos, 2016; Rehfeldt, Dillen, Ziomek, & Kowalchuk, 2007) and children with autism (Jackson, Mendoza, & Adams, 2014; Rehfeldt, Dillen, Ziomek, & Kowalchuk, 2007).

Most commonly, perspective shifting has been established using variants of the "Barnes-Holmes Protocol" (e.g., McHugh, Barnes-Holmes, & Barnes-Holmes, 2004; though see Guinther, 2017). The protocol incorporates a variety of self-focused distinction relations called **deictic relations**, which define the relation between the self and various aspects of the external world. Deictic relations have three dimensions. The I-YOU dimension distinguishes self from other people. The HERE-THERE dimension distinguishes between geographic locations of self versus others (HERE is always associated with I, and THERE always with YOU). The NOW-THEN dimension distinguishes between events taking place in the present versus other times.

The everyday importance of the deictic dimensions can be appreciated by considering the kinds of self-referencing verbal interactions that adults tend to have with young children. Adults may ask children questions like, "What is Mommy doing? (I/YOU), "Where are you?" (HERE/THERE), and "What did you do this morning at preschool?" (NOW/THEN). The reinforcing of correct answers forces attention to important deictic distinctions (e.g., it is difficult to correctly answer "Where are you?" without grasping that you are not your dog who is sitting in the other room). However, these are "simple" deictic relations that are not perspective shifting per se but rather prerequisites to it. Adults also ask children questions that require a reversal of perspective along one or more of the three deictic dimensions. For example, the question "What do you like to do when you are at Grandma's house?" requires a reversal of both HERE/THERE and NOW/THEN. This is an example of perspective shifting.

The Barnes-Holmes Protocol uses verbally presented scenarios about which the learner is asked to answer questions reflecting the three deictic dimensions. Table 20.2 shows examples of these scenarios. This verbal format differs from match-to-sample procedures that, in many cases, contain as little verbal content as possible.[17] The scenarios in the Barnes-Holmes Protocol represent three levels of deictic complexity (Table 20.2, leftmost column). *Simple* deictic relations involve the three core deictic dimensions mentioned earlier and, as mentioned previously, are not instances of perspective shifting. The building blocks of perspective shifting are *reversed* deictic relations (involving reversal along one of the three dimensions) and *double-reversed* relations (involving reversal on multiple dimensions).

You may be struck by how confusing some of the scenarios in Table 20.2 seem, even to an advanced learner like yourself. This illustrates that deictic relations do not necessarily "come naturally" and thus may not always be acquired under everyday circumstances (even by high-functioning adults; look ahead to Figure 20.11). It is therefore noteworthy that the Barnes-Holmes Protocol has been successfully used to teach perspective shifting to young children and persons with disabilities (e.g., Barron, Verkuylen, Belisle, Paliliunas, & Dixon, 2018; Belisle, Dixon, Stanley, Munoz, & Daar, 2016; Heagle & Rehfeldt, 2006; Jackson et al., 2014; Rehfeldt et al., 2007).

Intellectual Development

Intelligence

The construct of *intelligence* traces back to well before psychology was an independent discipline, and psychologists have sought to measure it for more than 100 years. Individual

TABLE 20.2 Example Items from the Barnes-Holmes Protocol for Teaching Perspective Shifting

LEVEL	*Example Item*[a]	I/YOU	HERE/THERE	NOW/THEN
Simple	I have a red brick and you have a green brick. Which brick do I have?	Simple		
	I am sitting here on the blue chair and you are sitting there on the black chair. Where are YOU sitting?	Simple	Simple	
	Yesterday I was watching television. Today I am reading. What am I doing NOW?	Simple		Simple
Reversed	I have a red brick and you have a green brick. If I was YOU and YOU were ME, which brick would I have?	Reversed		
	I am sitting here on the blue chair and you are sitting there on the black chair. If I was YOU and YOU were ME, where would YOU be sitting?	Reversed	Simple	
	I am sitting here on the blue chair and you are sitting there on the black chair. If HERE was THERE and THERE was HERE, where would YOU be sitting?	Simple	Reversed	
	Yesterday I was watching television. Today I am reading. If NOW was THEN and THEN was NOW, what would I be doing THEN?	Simple		Reversed
Double reversed	I am sitting here on the blue chair and you are sitting there on the black chair. If I WAS YOU and YOU were ME, and HERE was THERE and THERE was HERE, where would YOU be sitting?	Reversed	Reversed	Simple
	Yesterday you were sitting here on the blue chair and today you are sitting there on the black chair. If I was YOU and YOU were ME, and NOW was THEN and THEN was NOW, where would I be sitting?	Reversed	Simple	Reversed
	Yesterday I was sitting here on the blue chair and you were sitting there on the black chair. If HERE was THERE and THERE was HERE, and if NOW was THEN and THEN was NOW, where would I be sitting?	Simple	Reversed	Reversed

[a]Items can represent three different deictic dimensions (I/YOU, HERE/THERE, and NOW/THEN). Each dimension can be represented in non-deictic (simple) or deictic (reversed and double-reversed) form.

Based on Table 1 of "Teaching Perspective-Taking to Typically-Developing Children Through Derived Relational Responding," by A. I. Heagle and R. A. Rehfeldt, 2006, *Journal of Early and Intensive Behavioral Interventions*, *3*, 1–34.

differences in intelligence are seen as having practical importance; for instance, intelligence tests are used in schools for diagnosing learning issues and for placing children into educational tracks (Kaufman, 2018). Suffice it to say that any theoretical or empirical effort that sheds light on *intelligence* is likely to be of substantial interest.

The chapter introduction suggested that "knowledge" and "understanding" consist of derived stimulus relations (e.g., Figure 20.1). If this is true, then three things follow. First, people whose experiences create a wide variety of derived stimulus relations should be perceived as especially knowledgeable. Second, because what is already known is part of the raw material for new stimulus relations, those who "know a lot" have an advantage in future learning—they have more conceptual hooks, as it were, on which to hang new learning. Third, recall that rich histories of relational learning are the source of the generalized relational repertoires (AARR) that RFT hypothesizes. This may allow people to learn new things with unusual speed and flexibility, which is one basis on which they might be called *intelligent*. On the basis of these observations, RFT theorists have equated "general intelligence" with AARR (Barnes-Holmes, Barnes-Holmes, et al., 2001; see also Cassidy, Roche, & O'Hora, 2010), and have begun to study the connections between relational repertoires and performance on standardized measures of intelligence.

Early results are encouraging. Several investigations have found positive correlations between standardized intelligence test scores and performance on relational tasks, including those emphasizing hierarchical relations (Mulhern, Stewart, & McElwee, 2017), temporal relations (O'Hora et al., 2008), and distinction relations (O'Hora, Pelaez, & Barnes-Holmes, 2005), as well as a battery involving multiple relation types (Colbert, Dobutowitch, Roche, & Brophy, 2017). Even more exciting, several studies have shown that intensive relation-focused interventions can significantly increase children's standardized intelligence test scores (Amd & Roche, 2018; Cassidy, Roche, Colbert, Stewart, Grey, 2016; Hayes & Stewart, 2016; Parra & Ruiz, 2016). These outcomes suggest a potentially important link between the kind of behavior change that applied behavior analysts target and larger scale measures of functioning that are valued by schools and other institutions, making this an area ripe for further investigation.

Developmental Trends in Relational Behavior

On a related topic, if relational abilities are learned over time, as RFT proposes, then one would expect to see orderly age-correlated changes in these abilities. To be clear, we are not suggesting that relational abilities automatically develop as a function of age. Rather, age is a marker for multiple-exemplar relational experience that, under the right circumstances,

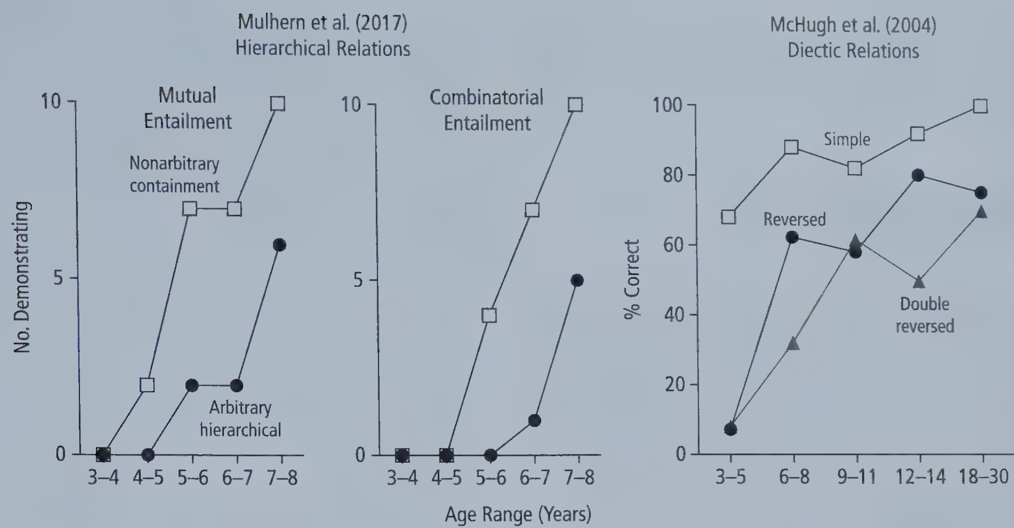


Figure 20.11 Left: Number of children (out of 10) in several age groups who demonstrated mutual and combinatorial entailment based on hierarchical stimulus relations. Each panel compares competence on a relatively simple (nonarbitrary containment) and relatively complex (arbitrary hierarchical) categorization. Right: Percentage of correct responses by children in several age groups who demonstrated self-referential relations. Results are shown separately for non-deictic (simple relations) and two kinds of deictic relations.

Left: Based on Table 5 of "Investigating Relational Framing of Categorization in Young Children," by T. Mulhern, I. Stewart, & J. McElwee, *Psychological Record*, *67*, 519–536. Right: Adapted from Figure 1 of "Perspective-taking as Relational Responding: A Developmental Profile," by L. McHugh, Y. Barnes-Holmes, & D. Barnes-Holmes, *Psychological Record*, *54*, 115–144.

accumulates over time. It is entirely possible for an individual of any age to lack the relevant experience to exhibit a particular type of generalized relating. A few investigations have directly sought evidence of the development of relational abilities. Two studies, each using a longitudinal design with a single child, found evidence of equivalence-class formation only after about 19 to 23 months of age (Lipkins, Hayes, & Hayes, 1993; Luciano et al., 2007). Such findings are interesting for two reasons. First, in these studies, equivalence-class formation was preceded by the appearance of presumably prerequisite relational abilities, such as symmetrical relations. Second, normatively speaking, it is around this same age that children first reliably exhibit verbal behavior (e.g., the naming of objects) that might be thought to draw upon equivalence relations (Luinge, Post, Wit, & Goorhuis-Brouwer, 2006).

Regarding how early in development relational abilities can arise, the longitudinal studies just described are somewhat ambiguous, because they provided relational experience along with assessing relational abilities. Other studies have used cross-sectional designs that provided only enough relational experience to assess performance on a single occasion, thereby reducing the chances that the research protocol instated any new abilities. Figure 20.11 (left panel) summarizes selected findings of a study that assessed hierarchical relations in children of different ages (Mulhern et al., 2017). The figure shows the number of children, out of 10, in each age cohort who met a mastery criterion for each of several kinds of relations. The general pattern was for derived mutual entailment and combinatorial entailment based on hierarchical classes to become more prevalent with age. Shown separately are the results for relatively difficult hierarchical relations versus nonarbitrary containment, an elementary, nonhierarchical type of categorization that is based on simple physical properties of stimuli like size and color. Developmental trajectories were steeper for nonarbitrary containment, which all children demonstrated in the age 7 to 8 cohort, than for hierarchical relations, which were not reliably present even for the older children who were tested. This is consistent with the general assumption that "difficult" levels of categorization develop relatively slowly (Bornstein & Arterberry, 2010).

Figure 20.11 (right panel) summarizes the findings of a study that assessed deictic relations in individuals aged 3 to 30 years (McHugh et al., 2004). For each relation type, the figure shows the percentage of assessment trials on which these individuals answered correctly. Simple deictic relations, which do not involve perspective shifting, were evident even in the youngest children. Reversed and double-reversed relations, which do involve perspective shifting, were absent before age 6, and not completely reliable even in the 18 to 30 age cohort. This finding is intriguing for two reasons. First, the spotty acquisition of reversed/double-reversed relations is consistent with the assumption that perspective shifting is a difficult skill—for instance, compare to the developmental pattern for hierarchical relations in the left panel. Second, the trend for reversed/double-reversed relations is broadly consistent with other research reporting that, in normally developing children, reliable evidence of self-focused constructs like "theory of mind" arises around age 5. Note, however, that there may be cultural differences in developmental trajectory, as Chinese children show "theory of mind"

earlier than U.S. children (Sabbagh, Xu, Carlson, Moses, & Lee, 2006). Assuming that "culture" implies different kinds of social learning environments, this finding is consistent with the RFT position that relational abilities are learned.

DERIVED STIMULUS RELATIONS AND GENERAL WELL-BEING

The previous section addressed some big-picture psychological constructs to which derived stimulus relations of various types may be related. Here is one more: general psychological well-being. "Well-being" can mean many things, but a critical component likely involves behaving in ways that make contact with one's most valued reinforcers. A three-term-contingency view assumes that behavior follows reinforcement, so a question of interest is when this would ever *not* happen. As we describe below, one answer intersects with what is known about derived stimulus relations, and this sets the stage for **acceptance and commitment therapy** (ACT), an evidence-based behavior therapy that counts RFT among its influences and is designed to help people better contact their reinforcers (Hayes, 2005; Hayes, Strosahl, & Wilson, 2016; see also McEnteggart, 2018).

Soon we will describe selected aspects of ACT interventions, which, unlike conventional ABA interventions, rely heavily on verbal exercises (e.g., Assaz, Roche, Kanter, & Oshiro, 2018).[18] This is different enough from the *modus operandi* of conventional ABA that we must begin by explaining ACT's conceptual undergirding so that its verbally based interventions make sense. Doing so will also clarify how ACT connects to the derived stimulus relations described here and in Chapter 19 (for more on these connections, see Assaz, et al., 2018; McEnteggart, 2018).

Because some effort will be required to digest ACT's precepts, let us try to provide some motivation for doing so: ACT works. Behavior analysts are interested in data and, as an exercise in reinforcer sampling, Figure 20.12 presents some, in the form of outcomes from an intervention for direct support staff who worked with clients with developmental disabilities (Chancey et al., in press). During baseline the staff engaged clients in positive interactions only rarely. After an intervention began, positive interactions with clients more than doubled. Note, however, that this intervention was not preceded by a functional assessment (Chapter 27) and included no prompts or consequences directly targeting interactions with clients. That is, the intervention was effective despite omitting much that is used in conventional ABA. Later we will describe the intervention, but for now we proceed to the present section's punchline: We think that applied behavior analysts should be interested in *any* intervention that works, and curious about *how* it works, even if that means stepping outside the confines of familiar practices. In reading about ACT, prepare yourself for a striking departure from the type of analysis employed in conventional ABA.

How Private Events Curtail Contact with Reinforcers

The core rationale of ACT is that private events (see Chapter 18) are at the root of a variety of human challenges (Hayes, 2005). The following is a trivial example of how private events can be problematic. You are waiting for a bus to take you to work, which is a significant source of reinforcers. A woman walks by who reminds you of an old friend you haven't seen since high school. You begin to think about experiences you and your friend shared and you try to imagine what she might be doing today. This reverie ends suddenly with the sound of your bus—which you hadn't noticed approaching or loading—pulling away. By being "in your mind" (Hayes, 2005) rather than in the situation, you have missed the bus, and with it a good part of your work day. This isolated incident illustrates how private behaviors can compete with public behaviors. Now imagine what it would be like to be constantly "in your mind." Engaging in an extensive suite of private behaviors might cause you to miss a lot of reinforcers that are contingent on your public behavior. This experience is common in a variety of clinical disorders (Chawla & Ostafin, 2007), and indeed to many people without clinical disorders who fail to get full enjoyment out of life (Hayes, 2005). Box 20.6 provides a fictional, but not unrealistic, example of the latter.

Unpleasant Private Events Monopolize Attention and Occasion Problem Public Behaviors

As Box 20.6 suggests, one way in which private events compete with "well-being behaviors"—those that contact high-priority reinforcers—is simply to crowd them out. Aversive events may evoke unpleasant emotional responses that interfere with well-being behaviors (see Friman, Hayes, & Wilson, 1998). This is similar to how, for laboratory animals, "fear" elicited by electric shocks can interrupt positively reinforced behavior (e.g., Estes & Skinner, 1941). Unpleasant emotions also can potentiate public behaviors that interfere with well-being behaviors. The woman who kicks her dog after accidentally striking her thumb with a hammer illustrates the phenomenon, as do individuals who become socially aggressive after experiencing unpleasant events (e.g., Roberton, Dafferm, & Bucks, 2012).[19] Also relevant is a host of possible avoidance behaviors. For instance, imagine that a college professor must write a lengthy book chapter on a tight time line. Through stimulus relations formed during past experiences involving missed deadlines, working on the chapter elicits considerable anxiety, which the professor escapes by taking long bicycle rides when he should be writing. In this strictly hypothetical example (having *nothing* whatsoever to do with either of the present authors), biking is negatively reinforced and occurs with increasing frequency, even though in the long run this compromises reinforcers associated with finishing the chapter.

Transformation of Function Increases Contact with Suffering

Unpleasant events like being mugged (Box 20.3) may happen only once, but through transformation of function in derived stimulus relations they may be experienced repeatedly. To illustrate, let's return to the school-anxious boy, Shawn, described in Box 20.1. The boy had an unpleasant moment with his teacher, but suffering related to that moment expanded to fill a good portion of his life away from the teacher. The "spread of suffering" (see Dymond et al., 2014) began when direct experience (the reprimand) transformed the function of one stimulus (the teacher) in an already-existing stimulus class (Figure 20.4).

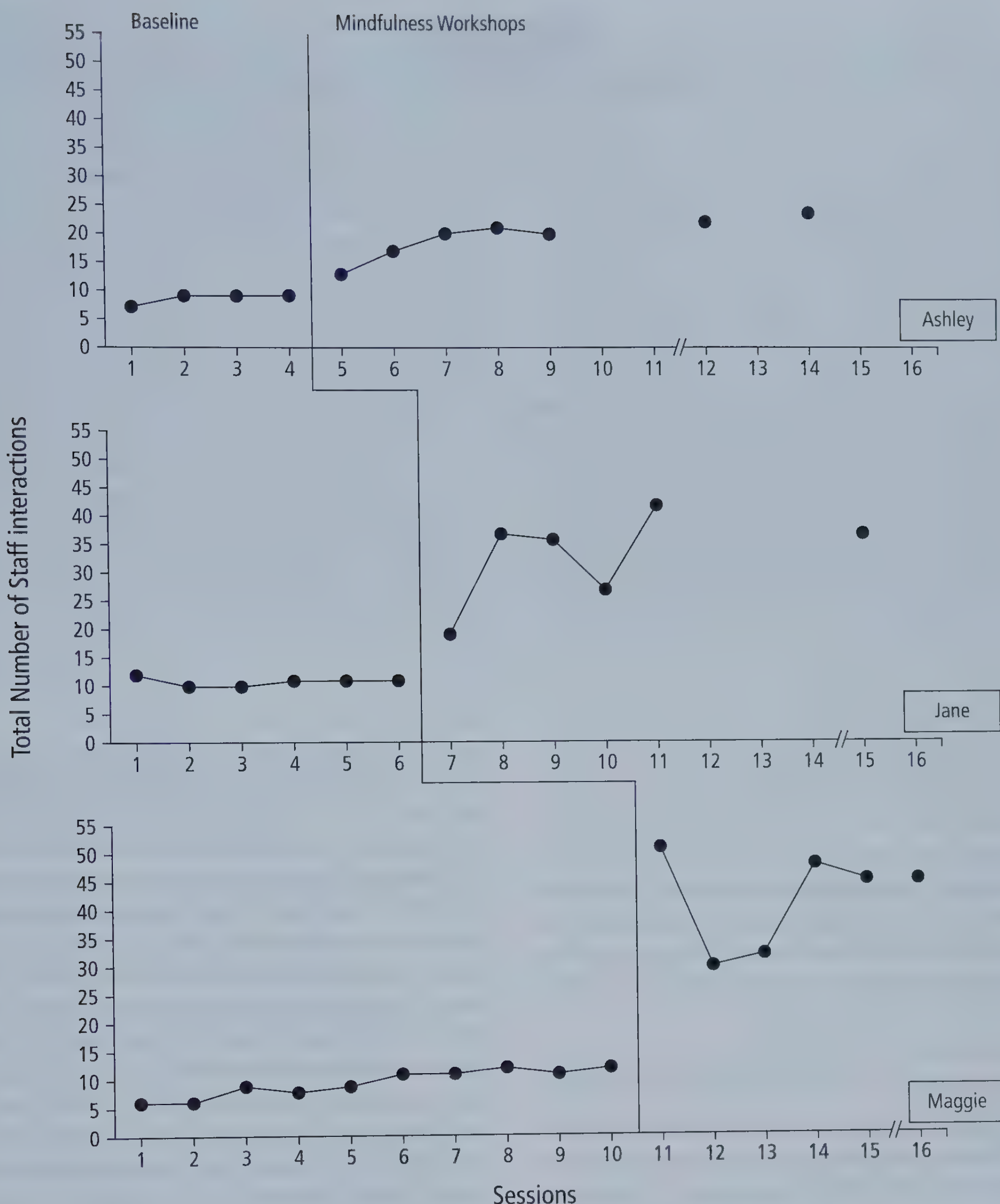


Figure 20.12 Effects of an intervention that reduced frequency of serious aggression by three adolescents.

Republished with permission of Elsevier, from The Impact of Mindfulness Skills on Interactions Between Direct Care Staff and Adults with Developmental Disabilities, C. Chancey, et al., in press, *Journal of Contextual Behavioral Science*. Permission conveyed through Copyright Clearance Center, Inc.

Other stimuli in the class acquired the same function. For Shawn, this functionally was like being yelled at by his teacher multiple times every day.

Less direct effects are possible. Let's say that on vacation you are having a lovely time visiting a Scottish castle. You adore the architecture, are impressed with the beauty of the medieval tapestries hanging on the walls, and find the history described by the tour guide to be fascinating. Thus, you're deriving an intensely pleasant stimulus class involving the castle and all of these stimuli. Then you stumble upon a coat of arms, which reminds you of your mother, who was descended from Scots and treasured the coat of arms that had been handed down through her family. Your mother passed away 2 years ago after a painful illness, and thinking of her makes you feel melancholy. This new transformation of function means that the stimulus class involving the castle will no longer involve untempered joy.

Finally, generalized relational abilities can contribute to the "spread of suffering" by transforming pleasant functions into unpleasant ones:

> Consider a six-year-old girl who drew a picture at school. Her father says, "That's nice. You did really well on that!" Given her verbal competence, the child will also be in contact with the opposite of what is said, that is, "not so nice" and "not doing very well." She may not contact this at the moment her father speaks, but she may do so at a later time, such as when she tries to draw another picture

BOX 20.6

The Private Events of Walter Mitty

> "We're going through!" The Commander's voice was like thin ice breaking. He wore his full-dress uniform, with the heavily braided white cap pulled down rakishly over one cold gray eye. "We can't make it, sir. It's spoiling for a hurricane, if you ask me." "I'm not asking you, Lieutenant Berg," said the Commander. "Throw on the power lights! Rev her up to 8,500! We're going through!" (Thurber, 1999, p. 55)

So begins the iconic James Thurber short story, "The Secret Life of Walter Mitty" (which was loosely adapted into a 2013 film starring Ben Stiller and Kristin Wiig). The story follows a middle-aged man through a series of afternoon errands in town with his wife. There is no real hurricane or Lieutenant Berg; these are part of elaborate fantasies in which Mitty engages while the day's events unfold around him. At various turns, Mitty pictures himself a brave military commander, a famous surgeon, and a convict facing a firing squad.

The story is usually said to illustrate a man's desperate efforts to escape from a boring life, but quite possibly his life is boring (devoid of external reinforcers) *because* of the fantasies. Each time Mitty lapses into a fantasy, to all the world he becomes behaviorally inert. He misses out on conversations. He fails to move his car forward when the traffic light changes. He forgets what to buy in a store. The private behaviors of his fantasies crowd out public behaviors that might contact reinforcers. As a result, other people—including his wife—think Mitty strange and either withdraw from him or browbeat him in attempt to prompt useful behavior. This magnifies Mitty's isolation.

> and it doesn't turn out the way she wanted it to. She will not need anyone else to make the judgment that this is "not so nice" or that she "did not do well." (Torneke, 2010, pp. 136–137)[20]

In the end, it doesn't matter whether unpleasant experiences are "real" or derived—in fact, there is no difference in the experience of the two (Friman et al., 1998). Because of transformation of function within stimulus classes, unpleasant private events that constrain well-being behaviors can occur any place, at any time.

Excessive Rule Following Interferes with Well-being Behaviors

Recall the learned helplessness paradigm described in Box 20.3. Unpleasant experiences, like being mugged, can lead to the creation of generalized self-rules that occasion behaviors that interfere with well-being behaviors (recall how an "I'm ineffective" rule undermined college classes and social relationships). **Rule-governed behavior**, also called *instructional control*, occurs when verbal stimuli (*rules*) occasion behavior, alter motivating operations, and/or specify the consequences of that behavior (Hayes, 1989; Skinner, 1969). Rules and instructions can come from other people, as when a parent tells a child, "To avoid being hit by a car, look both ways before crossing the street." However, people contacting environmental contingencies also generate their own rules, and these verbal stimuli can control behavior just as externally generated rules do (e.g., Rosenfarb, Newland, Brannon, & Howey, 1992). Behavior under the control of rules (both self- and other-generated) often is insensitive to change by local contingencies (Hayes, 1989; Skinner, 1969).

RFT views rule governance as operating through derived stimulus relations. Rules are thought to control behavior partly because they include stimuli (words) that participate in stimulus classes with "real" (nonverbal) stimuli that control behavior. For example, for a student whose failed geometry test evokes the rule "I am terrible at math," the words "terrible" and "math" are experienced in much the same way as the actual unpleasant experience (failing the test) to which they are related. You might recognize self-rules as similar to deictic relations, person-location-time relations in the format of I-HERE-NOW. Self-rules can pervade everyday experience and broadly disrupt well-being behaviors when they are extended along deictic dimensions. The self-rule "I am terrible at math" is based on an experience that occurred at a specific time and place, but describes (hypothesized) outcomes at all times and places. This is an extension on the HERE/THERE and NOW/THEN dimensions. The same happened with the mugging victim of Box 20.3: A factual statement (*I* was not effective *in that place at that time*) extended to other situations where it might not be accurate but still could control behavior (*I* am ineffective *at all times* and *in all places*). Other problem rules may conflate the I/YOU terms, as when a man whose father was divorced 4 times derives that "*I* can never find lasting happiness in a relationship" (it is the father, not the son, whose relationship skills are in question). For an accessible general introduction to how stimulus relations facilitate rule governance, see Torneke (2010, Chapter 6).

Self-rules become problematic when the public behaviors they control interfere with well-being behaviors. Box 20.7 describes an example. Note especially in this example how behaviors controlled by self-rules are maintained by negative reinforcement. In cases like this, unpleasant private events create the motivating operations for both public behavior and the self-rules that control it. Behavior that terminates these private events is automatically strengthened (e.g., Dymond & Roche, 2009). Thus, for the school-avoidant boy in Box 20.1, hiding in the neighborhood until after the school bus departs prevents school-related fear. For the germ-obsessed parent in Box 20.7, cleaning reduces anxiety. This is one reason why rule-governed behavior is insensitive to change: It generates its own immediate reinforcers (McEnteggart, 2018).

BOX 20.7

The Sanitation Complication[21]

Joachim was the father of two girls, ages 2 and 3. When the first daughter was born, he and his wife decided that Joachim would be a stay-at-home dad, because Joachim's wife earned more money than he did as a first-grade teacher, and in that job Joachim had been great with kids. Joachim looked forward to teaching his girls about the world through "field trips," like those he used to arrange for his students, to the park, to museums, and to the zoo. He was excited to help the girls develop strong basic academic skills before they entered school.

But Joachim struggled unexpectedly in his role as full-time parent. For reasons even he did not understand, Joachim worried constantly about protecting his daughters from germs, and he developed a suite of self-rules like "I have to keep my children safe from disease in order to be a good parent," and "I need to keep the house clean." Before long Joachim became a cleaning machine. He walked around the house with a bottle of cleaning spray, wiping down every doorknob and light switch that he passed. He scrubbed the floors and disinfected every surface of the kitchen multiple times each day.

These behaviors "worked" in the sense that they left the home spotless and reduced Joachim's anxiety about germs. But they were incompatible with many of the things Joachim had planned to do as a parent. As a result, he rarely played with the girls, and when he did, the games were interrupted by frequent spot cleaning. Joachim did not arrange any "field trips" or teach the girls to read because he was too busy disinfecting. In short, Joachim became an absentee parent in his own home. This had adverse effects on his daughters, and it deprived Joachim of many of the joys of interacting with them and of helping them develop as people. At a conscious level Joachim understood this tradeoff, but the notion of being responsible for a dangerously unsanitary environment was so aversive that he simply did not think he could help himself.

How ACT's Verbally Guided Exercises Aim to Short-circuit Harmful Private Events

In ACT, the persistent interference of private events with well-being behaviors is called *psychological inflexibility* (Hayes, 2005). We will call this **behavioral inflexibility**, and it is the opposite of *flexibility*, a capacity to be "in the moment" such that one's behavior is under the real-time stimulus control of events happening in the current situation (Hayes, 2005). ACT interventions use mainly verbally guided exercises to disrupt inflexibility and to potentiate behavior that contributes to well-being (Harris, 2008; Hayes et al., 2016). One type of ACT exercise, called *values clarification*, is designed to help people identify their most valued positive reinforcers in important life domains such as relationships, work, and leisure (Lundgren, Dahl, & Hayes, 2008). Among other functions, these exercises aid in recognizing behaviors on which the high-priority reinforcers are contingent (and with which private events may interfere).

A second kind of ACT exercise is designed to alter the aversive function of private events. Some exercises mimic exposure therapies (e.g., Wolpe, 1961) by forcing repeated contact with anxiety-provoking thoughts—for example, by repeatedly speaking aloud those thoughts until their emotional properties abate and the thought-statement is experienced merely as sound; or by imagining the anxiety-provoking thought being spoken in the ridiculous voice of a cartoon character (Harris, 2008).

A third kind of ACT exercise is designed to lessen control of private events over other behavior. An organizing concept is *mindfulness*, which has been defined as "bringing one's complete attention to the present experience on a moment-to-moment basis" (Marlatt & Kristeller, 1999, p. 68), and as observation without judgment (Baer, 2003). Some mindfulness exercises, like "Take Ten Breaths" in Box 20.8, focus on mundane external stimuli to distract attention from troublesome private events.[22] Others (like "Floating Leaves" and "Soles of the Feet" in Box 20.8) first focus attention on the mundane and then shift it to private thoughts and rules, including those that may be problematic. Here the goal is to teach how to notice private events without acting on them. For example, school-fearful Shawn (of Box 20.1) may feel anxious and afraid when he hears his school bus, but he can learn to observe these private events and still get on the bus. The cleaning-obsessed parent can notice that he repeats his rules about protecting his children without having to scrub floors every hour.

Effectiveness Evidence

By now it should be obvious that ACT interventions take a very different form than those of conventional ABA. We understand that readers who are accustomed to functionally analyzing behavior in terms of its relation to external stimuli and consequences may find ACT difficult to understand[23] and may, like some other behavior analysts, resist the suggestion that public behavior should be examined through the lens of private events (e.g., Baum, 2011; Fisher, Groff, & Roane, 2011). Yet, as empiricists, behavior analysts are committed to following the data, and there is considerable evidence supporting the effectiveness of ACT for "well-being" challenges in many domains. This includes mental health issues in schools (Burckhardt, Manicavasagar, Batterham, Hadzi-Pavlovic, & Shand, 2017; Van der Gucht et al., 2017) and in a variety of psychological conditions like anxiety, depression, addiction, chronic pain, and somatic health problems (e.g., A-Tjak et al., 2015; Hann & McCracken, 2014; Kanter, Baruch, & Gaynor, 2006). Positive effects are not, however, limited to clinical conditions. For example, ACT (or an intervention using some of its components) has reduced workplace stress (Ly, Asplund, & Andersson, 2014), increased college students' engagement with physical activity (Butryn, Forman, Hoffman, Shaw, & Juarascio, 2011), and improved athletes' performance in competition (e.g., Schwanhausser, 2009).

BOX 20.8

Three ACT Mindfulness Exercises

These exercises instruct "observation without judgment" (attending to stimuli without otherwise reacting to them). The goal of practicing them is not to prevent problematic private events but rather to disrupt their control over other behavior.

Take Ten Breaths (Harris, 2008, p. 35)

1. Throughout the day, pause for a moment and take ten slow, deep breaths. Focus on breathing out as slowly as possible, until the lungs are completely empty, and breathing in using your diaphragm.
2. Notice the sensations of your lungs emptying and your ribcage falling as you breathe out. Notice the rising and falling of your abdomen.
3. Notice what thoughts are passing through your mind. Notice what feelings are passing through your body.
4. Observe those thoughts and feelings without judging them as good or bad, and without trying to change them, avoid them, or hold on to them. Simply observe them.
5. Notice what it's like to observe those thoughts and feelings with an attitude of acceptance.

Floating Leaves on a Moving Stream (Hayes, 2005, pp. 76–77)

Imagine a beautiful, slow-moving stream. The water flows over rocks, around trees, descends downhill, and travels through a valley. Once in a while, a big leaf drops into the stream and floats away down the river. Imagine you are sitting beside that stream on a warm, sunny day, watching the leaves float by.

Now become conscious of your thoughts. Each time a thought pops into your head, imagine that it is written on one of those leaves. If you think in words, put them on the leaf as words. If you think in images, put them on the leaf as an image. The goal is to stay beside the stream and allow the leaves on the stream to keep flowing by. Don't try to make the stream go faster or slower; don't try to change what shows up on the leaves in any way. If the leaves disappear, or if you mentally go somewhere else, or if you find that you are in the stream or on a leaf, just stop and notice that this happened. File that knowledge away and then once again return to the stream, watch a thought come into your mind, write it on a leaf and let the leaf float away down the stream.

Meditation on the Soles of the Feet (Singh et al., 2011, p. 1055)

This exercise was used in an intervention, to be discussed shortly, to reduce aggression by adolescent boys (results are summarized in Figure 20.12).

1. If you are standing, stand in a natural rather than an aggressive posture, with the soles of your feet flat on the floor.
2. If you are sitting, sit comfortably with the soles of your feet flat on the floor.
3. Breathe naturally, and do nothing.
4. Cast your mind back to an incident that made you very angry. Stay with the anger.
5. You are feeling angry, and angry thoughts are flowing through your mind. Let them flow naturally, without restriction. Stay with the anger. Your body may show signs of anger (e.g., rapid breathing).
6. Now, shift all your attention fully to the soles of your feet.
7. Slowly, move your toes, feel your shoes covering your feet, feel the texture of your socks, the curve of your arch, and the heels of your feet against the back of your shoes. If you do not have shoes on, feel the floor or carpet with the soles of your feet.
8. Keep breathing naturally and focus on the soles of your feet until you feel calm.
9. Practice this mindfulness exercise until you can use it wherever you are and whenever an incident occurs that may otherwise lead to you being verbally or physically aggressive.
10. Remember that once you are calm, you can walk away from the incident or situation with a smile on your face because you controlled your anger. Alternatively, if you need to, you can respond to the incident or situation with a calm and clear mind without verbal threats or physical aggression.

Interventions using ACT or some of its components have reduced the stress and depressive symptoms of parents of children with autism spectrum disorder (ASD; Cachia, Anderson, & Moore, 2016; Singh et al., 2014), as well as developmental disabilities staff (Biglan, Layton, Jones, Hankins, & Rusby, 2013; McConachie, Mckenzie, Morris, & Walley, 2014). ACT also has been used to increase positive interactions between staff and clients (Castro, Rehfeldt, & Root, 2016). Figure 20.12, which we introduced earlier with little explanation, illustrates with data from a study by Chancey et al. (in press). Staff completed mindfulness exercises, but the intervention did not focus on client interactions per se, and there was no *in situ* prompting or reinforcement of interactions with clients. Rather, improved interactions presumably were a diffuse side effect of improved staff well-being.

Because of ACT's reliance on verbal exercises, it is not yet known how broadly applicable it may be to persons with intellectual disabilities. However, a few published applications exist, all focusing on higher-functioning individuals. In one case, ACT exercises targeting anxiety, coupled with behavioral skills training, improved job-interviewing performance in two of three individuals with developmental disabilities (Brazeau et al., 2017). A randomized controlled trial found mindfulness exercises to reduce symptoms of depression and anxiety in

adults with an ASD diagnosis (Spek, Van Ham, & Nyklicek, 2013). For a further example, examine Figure 20.13, which shows the effects of mindfulness practice on aggression in three adolescents with an ASD diagnosis (Singh et al., 2011). In this case, the mindfulness exercise ("Soles of the Feet" in Box 20.8) was expressly intended to create replacement behavior for aggression. During the first week of the intervention, the boys practiced the exercise with parental assistance for 30 minutes. Afterward they practiced under parental supervision twice per day until aggression was absent for 4 consecutive weeks. Practice stopped at that point, but the boys retained a digital audio recording of the exercise instructions that could be used as desired. Serious aggression (hitting, kicking, and biting), which occurred during baseline about 2 to 3 times per week, dropped to near-zero levels following introduction of mindfulness exercises and stayed there for about 3 years.

Although few published ACT interventions have targeted persons with developmental disabilities directly, an especially interesting phenomenon is that when ACT is used to improve the well-being of parents or staff, those they care for may experience positive trickle-down effects (Cachia et al., 2016). Figure 20.14 shows an example (Singh et al., 2014). Participants were three mothers of high-functioning adolescents with an ASD diagnosis. All three had in the past worked with applied behavior analysts but had abandoned behavioral programs because of the stress of daily implementation; they had not used behavioral programs for at least 2 years prior to the study. Across 8 weeks, the mothers completed a once-weekly workshop focusing on mindfulness-based meditation. Afterward, they were instructed to continue practicing mindfulness each morning, and parent diaries showed that on most days they did. Mother stress, as measured by a self-report rating scale, decreased from baseline to the practice phase. Concomitantly, although the workshops did not focus on behavioral interventions or even on the changing of child behavior, incidents of child aggression decreased. Disruptive behavior also decreased, and child compliance with mother requests increased (not shown in Figure 20.14).

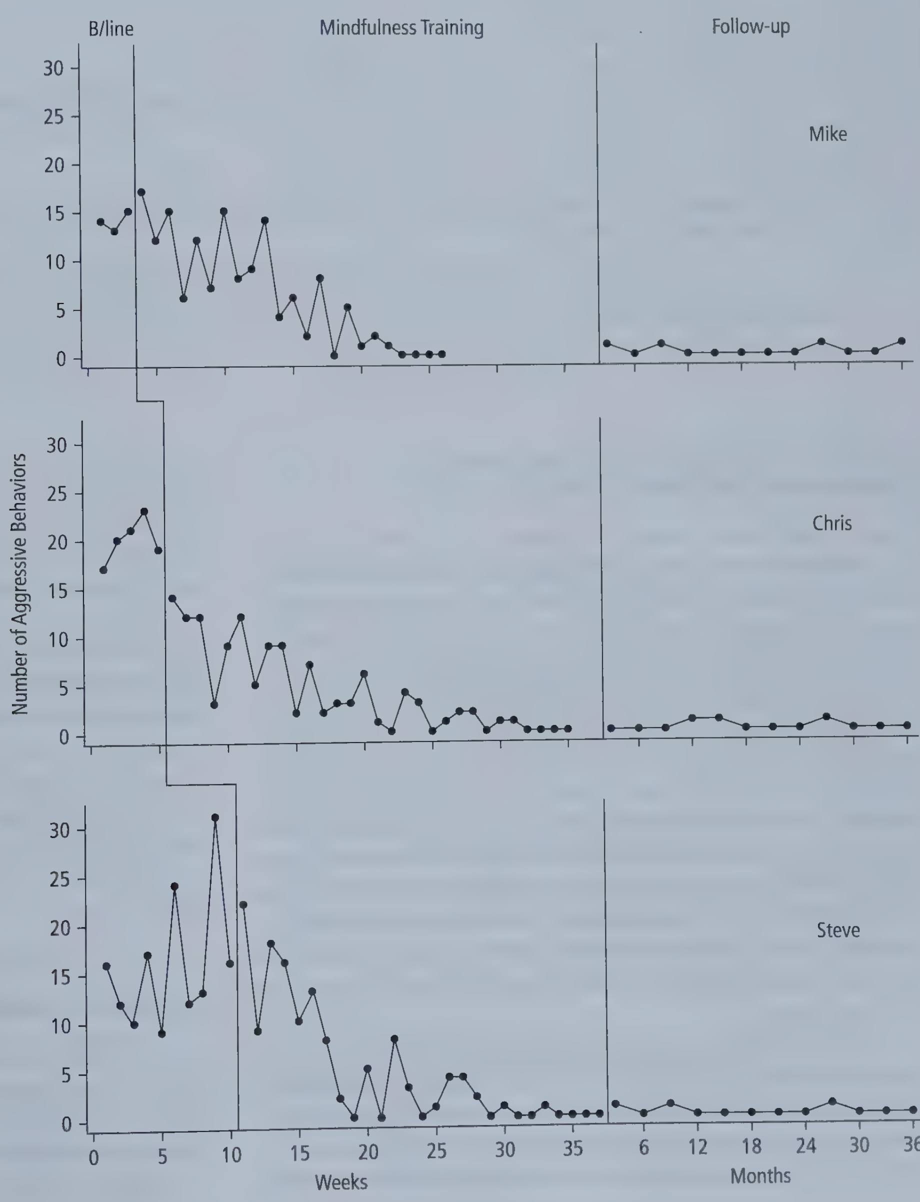


Figure 20.13 Effects of mindfulness practice on weekly incidents of physical aggression by three adolescents with autism.

Republished with permission of Elsevier, from A Mindfulness-Based Strategy for Self-Management of Aggressive Behavior in Adolescents with Autism, N. N. Singh and colleagues, 2011, *Research in Autism Spectrum Disorders*, *5*, pp. 1153–1158. Permission conveyed through Copyright Clearance Center, Inc.

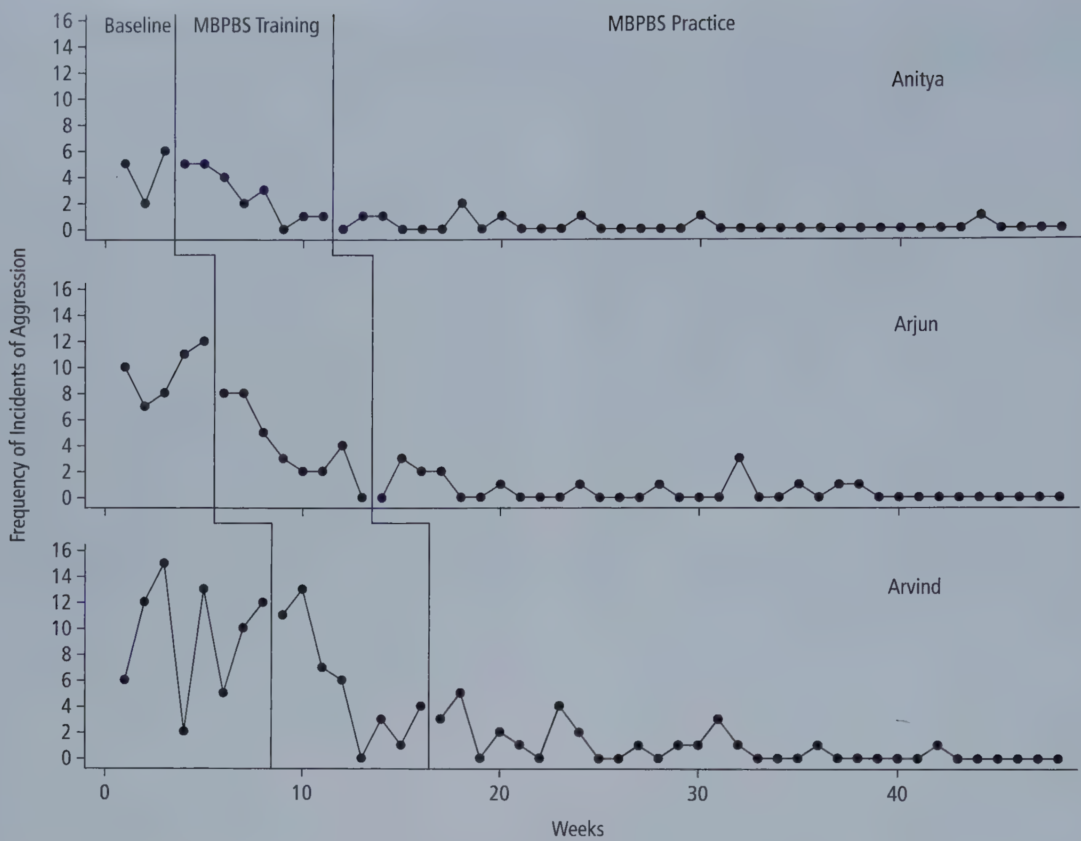


Figure 20.14 Weekly incidents of aggression by three adolescents with an autism spectrum disorder diagnosis.

Republished with permission of Springer, Mindfulness-based Positive Behavior Support (MBPBS) for mothers of adolescents with autism spectrum disorder: Effects on adolescent behavior and parental stress, N. N. Singh and colleagues, 2014, *Mindfulness*, *5*, pp. 646–657. Permission conveyed through Copyright Clearance Center, Inc.

Section Conclusion: The Functional Significance of "Well-Being"

Results like those shown in Figures 20.12 through 20.14 capture the essence of what is meant by "well-being" and of why applied behavior analysts should be interested in this construct. In contrast to most ABA interventions, those grounded in ACT do not usually target discrete problem behaviors, incorporate direct prompts, or systematically manipulate externally arranged consequences. They may teach no specific replacement behaviors. Yet by addressing very general aspects of functioning, they can have a positive impact on the kinds of problems that applied behavior analysts often address. Practitioners of ACT might argue, therefore, that conventional ABA has ignored an important context for behavior problems, that taking place within the individual. Well-being, from the perspective of ACT, exists when troublesome private events do not occasion problem behaviors or constrain behavior that contacts reinforcement.

Whether that argument registers or not, ACT is worth exploring for at least two reasons. First, relevant to the mission of the present chapter, ACT extends concepts of derived stimulus relations in unexpected directions, and this necessarily challenges the interested observer to make sure those concepts are well understood. Second, ACT illustrates how the scope of practice of an applied science can be much broader than what is often seen in ABA, which has remained heavily focused on persons with disabilities. Those who think ABA should systematically address the full range of human problems and interests (e.g., Friman, 2006; see also Chapter 1) should be intrigued by ACT's primary focus on typically developing people, and its contention that verbal behavior contributes to many human problems and can therefore be used as a tool of intervention. Those who are not persuaded by ACT's theoretical framework must reckon with data showing that it works and generate a compelling alternative explanation for why. Overall, we believe that anything that forces careful thinking about how to do the most good for the most people will be good for ABA and the people it serves.

A FINAL COMMENT

You would not be alone if you found reading the present chapter to be hard work. It is difficult to learn about equivalence classes (Chapter 19), and things become more complicated once the many additional ways of relating are taken into account. A goal of this chapter has been to persuade you that the effort

required to master this difficult topic is worth it. We have mentioned a number of ways in which derived stimulus relations concepts map onto important everyday behaviors and suggested some avenues for application. Yet even though research on derived stimulus relations began almost 50 years ago, bench-to-bedside translation has been fairly slow in developing (e.g., Critchfield, 2018; Dixon, Belisle, Rehfeldt, & Root, 2018; Rehfeldt, 2011).

There are good reasons for this. To learn about derived stimulus relations, one must first master a new vocabulary, although, as you have seen, derived stimulus relations can be complex enough that mere words do not do them justice (hence a heavy reliance on diagrams). This is a problem because most people don't think naturally in terms of the alternatives, like arrow-laden schematics. Humans, after all, are storytellers by nature (e.g., Hineline, 2018), not diagrammers and logicians. Something that this chapter and Chapter 19 can suggest but not fully illuminate is that, because derived stimulus relations have many moving parts, interventions focusing on them can be effortful to devise. No matter what potential may exist to harness derived stimulus relations concepts to create valuable behavior change, this can only happen after someone has carefully determined what to teach, how to teach it, and what derived abilities are expected to result from this.

The challenge of fully harnessing the potential of derived stimulus relations interventions now falls to you, the student, and your peers. To accomplish this, you must become so fluent with derived stimulus relations concepts, and the procedures arising from them, that applications and extensions suggest themselves effortlessly. Fortunately, there is ample precedent for mastering such challenges. Those who teach behavior analysis know that many people find even its most rudimentary concepts, like positive/negative reinforcement/punishment, to be extremely difficult. But as the present book illustrates, for those who persist, new concepts eventually become intuitive and essential building blocks for very sophisticated procedures.

Many opportunities exist to develop new derived stimulus relations interventions that target specific skills. Built properly, these are not just effective but also efficient, establishing new behaviors without directly teaching them (Critchfield, 2018; Rehfeldt, 2011; Chapter 19). In this regard, you might say that derived stimulus relations interventions can help ABA become better at what it historically has done. However, in this chapter we mentioned generalized relational repertoires (AARR), and devoted substantial space to big-picture constructs like intelligence and well-being, to illustrate that the stakes are higher than this. Looking past the complexities of the underlying concepts, you should have noticed, here and in Chapter 19, that derived stimulus relations research is not limited to particular populations or behavior problems. In this regard, it addresses the long-standing aspiration for behavior analysis to be a general-purpose science relevant to all of human experience (e.g., Skinner, 1953). There is one more consideration. Mastering derived stimulus relations concepts and following your curiosity about them may lead to work that (as the example of ACT illustrates) looks very different from conventional ABA. You may be concerned, therefore, about straying too far from the rich traditions that have made ABA so successful. To a large extent, derived stimulus relations concepts expand, rather than invalidate, ABA's foundations, but, inevitably, mastering and applying those concepts requires updates in how ABA does business. Fear not; we have it on good authority that change is acceptable.

> Regard no practice as immutable. Change and be ready to change again. Accept no eternal verity. Experiment. (Skinner, 1979, p. 346)

SUMMARY

What Are Nonequivalence Relations? Why Do They Matter?

1. *Stimulus relations* are arbitrary associations between stimuli, with "association" meaning that how one stimulus is responded to is intertwined in some way with how other stimuli are responded to.
2. Stimulus equivalence is only one of a wide variety of stimulus relations. Not all stimuli are related via "same as," and nonequivalence relations are integral to many everyday experiences.

The Vocabulary of Nonequivalence Relations

3. Key ways of talking about stimulus equivalence do not work for nonequivalence relations, so a substitute vocabulary is needed.
4. What's called *symmetry* in the stimulus equivalence literature is subsumed under the term *mutual entailment*, which means the converse of any A-to-B relation. For example, if the original relation is "A is greater than B," the converse will be "B is less than A."
5. What's called *transitivity* in the stimulus equivalence literature is subsumed under the term *combinatorial entailment*, which means any derived relation in which the function of stimuli depends on the other stimuli to which they are indirectly related.
6. What's called *transfer of function* in the stimulus equivalence literature is subsumed under the term *transformation of function*. This means that the stimuli in a relational class do not all acquire the same behavioral functions; however, the functions acquired by any stimulus depend on the functions acquired by others in the class.

Some Types of Nonequivalence Relations

7. Structured protocols similar to those used in equivalence-based instruction have been successful in establishing nonequivalence repertoires of everyday importance.

8. The notion that there are many ways of relating the same stimuli implies the existence of *contextual stimuli* that signal which way of relating is appropriate to a given situation. Explicit contextual stimuli are recommended as part of applied interventions focusing on derived stimulus relations.

9. *Distinction relations* involve relative stimuli being judged as different. *In opposition relations*, stimuli are opposites. In *comparison relations*, stimuli differ along some definable dimension.

10. Stimuli can also be related in terms of spatial or temporal juxtaposition and in terms of cause-effect relations.

11. Structured tools are just beginning to become available for programming derived stimulus relations like those described here.

Theoretical Foundations

12. Most research on nonequivalence relations has been guided by *relational frame theory*, which assumes (a) that verbal processes are bound up in derived stimulus relations, and (b) that given enough reinforced experience with multiple exemplars, people develop a higher-order ability to derive relations among sets of stimuli, even in the absence of further reinforced experience.

13. The role of verbal behavior in derived stimulus relations is not fully understood, but assuming a close connection has led to interesting research. For example, as the theory would predict, providing reinforced multiple-exemplar experience has been shown to create generalizable relational abilities. This may be an example of what relational frame theory calls *arbitrarily applicable relational responding*, a generalized repertoire that allows new classes to form without new reinforced experience.

14. It has also been found that, for some learners, the baseline relations of stimulus classes can be built simply by telling learners about them. Among other things, this finding could imply that verbal experience is functionally similar to the largely nonverbal match-to-sample experience provided in many stimulus equivalence protocols.

Nonequivalence Relations and Big-Picture Psychological Constructs

15. Categorical knowledge: Establishing a type of nonequivalence relation called *hierarchical relations* produces several effects seen in cognitive psychology studies of categorization.

16. Self-concept: A type of nonequivalence relation called *deictic relations* may underlie important self-related and social repertoires, a primary example of which is perspective shifting, in which an individual experiences the world as another person would.

17. Intellectual development: Normatively speaking, various types of derived stimulus relations tend to emerge gradually during child development. Importantly, systematically training stimulus relations has been found to boost scores on standardized intelligence tests.

Derived Stimulus Relations and General Well-Being

18. "Well-being behaviors" are those on which a person's most valued reinforcers are contingent. Many challenges to general psychological well-being involve the constraining of well-being behaviors by private events that include emotional responses and self-generated rules.

19. *Acceptance and commitment therapy* is a behavior therapy based partly on relational frame theory and research on derived stimulus relations (including nonequivalence relations). It focuses on promoting general well-being. In many cases, the therapy does not target public behaviors relevant to the referring problem. Instead, to undermine the control of behavior by problematic private events, the therapy employs verbally guided exercises that are very different from most ABA interventions.

20. The therapy has been shown to lessen symptoms of certain clinical disorders and to enhance performance in nonclinical situations. Another important application has been to reduce stress and alleviate depression in parents and caregivers of persons with disabilities. In some documented cases, improving caregiver well-being has been followed by improvements in client behavior.

KEY TERMS

acceptance and commitment therapy
arbitrarily applicable relational responding
arbitrary relations
behavioral inflexibility
causal relations
contextual stimulus
combinatorial entailment
deictic relations
derived relations
distinction relations
hierarchical relations
multiple-exemplar training
mutual entailment
nonequivalence relations
perspective shifting
relational frame theory
relational frame
rule-governed behavior
spatial relations
transformation of function
temporal relations

MULTIPLE-CHOICE QUESTIONS

1. Relational responding that experience does not directly teach, but that arises as an indirect function of experience is called:
   a. Emergent functions
   b. Derived functions
   c. Generative effect
   d. All of the above
   Hint: (See "The Vocabulary of Nonequivalence Relations")
2. The act of responding as if you see the world as from another's vantage point, or perceive a situation as if you were in another's position is called:
   a. Perspective shifting
   b. Empathy
   c. Social success
   d. Positive regard
   Hint: (See "Self Concept: Diectic Relations")
3. The core rationale of acceptance and commitment therapy suggests that the root of human challenges is/are:
   a. Private events
   b. Personality
   c. Past behavior
   d. All of the above
   Hint: (See "How Private Events Curtail Contact with Reinforcers")
4. In acceptance and commitment therapy, the persistent interference of private events with well-being behaviors is called:
   a. Maladaptive behavior
   b. Psychological inflexibility
   c. Psychosocial dysfunction
   d. Negative emotionality
   Hint: (See "How ACT's verbally guided Exercises aim to short-circuit harmful private events")
5. Classifying a stimulus as a member of a higher-order class on the basis that it is a member of the subclass that is a member of that higher-order class is called:
   a. Asymmetrical class containment
   b. Unilateral property induction
   c. Transitive class containment
   d. None of the above
   Hint: (See "Three characteristics of Hierarchical Categories")
6. The repeated acquisition of some behavioral repertoire with variations in stimuli, situations, or response feature is called
   a. Stimulus generalization
   b. Response generalization
   c. Single-exemplar training
   d. Multiple-exemplar training
   Hint: (See "Relational Frame Theory")

ESSAY-TYPE QUESTIONS

1. Discuss the types of nonequivalence relations.
   Hint: (See "Some Types of Nonequivalence Relations")
2. Name any two tools used for planning interventions involving nonequivalence relations.
   Hint: (See "Three Practical Considerations")
3. Discuss ACT's verbally guided exercises that aim to disrupt harmful private events.
   Hint: (See "How ACT's Verbally Guided Exercises Aim to Short-circuit Harmful Private Events")

NOTES

1. In keeping with tradition, we'll use the term *stimulus relations* as shorthand for various "ways of relating" (Reese, 1968; see also Critchfield, Barnes-Holmes, and Dougher, 2018). But the term is a bit of a misnomer. A stimulus relation is really a behavioral relation, a pattern in how an individual responds to groups of stimuli.
2. Compared to the semantic networks that cognitive researchers have mapped out, Figure 20.1 is quite rudimentary. *Beer* is related to many "bits of the world" that are not shown here, and each "bit" can itself be related to many other things. *Cold* is opposite to *hot*. *Lager* is made using *bottom fermentation*. A *baseball stadium* may remind you of your *father*, who used to take you to games, and so on.
3. But note that stimulus relations allow you to respond to a stimulus even when it is not physically present. For instance, you do not need to have a cold malt beverage in front of you to talk about *beer*, to picture how it looks, to imagine its taste, or to recall the embarrassing things you did last night after consuming too many. These responses can, instead, be occasioned by stimuli that participate with *beer* in arbitrary stimulus relations. This point is especially important to the chapter's final section.
4. Each trial was similar in format to a multiple-choice question. For example, when shown a picture of the frontal lobe, the student's job was to select an answer (*impulsiveness*) to which it was related, and ignore answers to which it was not (*impaired vision*, *spatial perception problems*, and *hearing problems*). Chapter 19 called this *selection-based* responding, and it likely allowed the students to succeed by responding on the basis of "any relation that exists," not of equivalence specifically.
5. The third derived property of equivalence classes, reflexivity, will not be addressed here. Reflexivity implies that a stimulus in an equivalence class is recognized as itself. This is an assumption for nonequivalence classes as well.

6. The word *entailment* means simply that one thing is dependent on another, so *mutual* entailment is when one direction of a relation (A to B) depends on, or is predictable from, the other (B to A).
7. This description is our contrivance for ease of explanation. Actually, the stimuli were several colored disks of paper, arrayed on a table surface such that the leftmost or rightmost disk served as a symbol of "many" or "few." Technically, then, this procedure confounded spatial relations (described below) with opposition relations, although that does not matter for present purposes.
8. This approach appears very different from the match-to-sample protocols described in Chapter 19. We will say more about the difference in the "Theoretical Foundations" section.
9. To repeat a point from earlier in the chapter, this might be construed as higher-order pairing, in which a neutral stimulus is paired with a stimulus that has been paired with an already-effective consequence. However, because higher-order conditioning typically is weak, it is unclear whether this explanation can account for effects like those just described.
10. There is an additional problem: For reasons too complicated to deal with here, generalized relational repertoires (arbitrarily applicable relational responding) are unlikely to arise from incoherent stimulus relations, even with multiple exemplars.
11. This phenomenon, of nonidentical stimuli being "the same" by virtue of inclusion in a common category, is one basis for *hierarchical* relations that we discuss later in the chapter.
12. Except in a handful of studies in which contextual cues are the explicit topic of investigation (e.g., Bush et al., 1989).
13. Naming theory (Horne & Lowe, 1996), introduced in Chapter 19, makes a similar assumption.
14. These studies illustrate a possible reason why "telling" and "demonstrating" procedures, which require no overt practice and provide no feedback, can produce derived stimulus relations—they may tap into already-established AARR.
15. What's also unclear is just how broadly applicable these generalized repertoires are expected to be. The RFT literature is somewhat vague on this point. The crux of the matter may concern what, precisely, functions as a contextual stimulus. The RFT literature tends to refer to discrete stimuli (like those verbal labels such as "opposite"), but if contextual stimuli can be compounds of discrete stimuli plus aspects of the setting where conditioning took place, as suggested earlier, then the repertoires they control might *not* readily generalize from one setting to another.
16. The traditional term is perspective *taking*, but that term normally is used to emphasize *only* the difference in perspective between self and other. We actually shift perspectives along multiple dimensions, as elaborated in this section.
17. Although the Barnes-Holmes Protocol has been used with children as young as 3 years of age, because of its narrative format it may yield the best results with learners at a developmental age of about 5 or older (e.g., Montoya-Rodriguez, Molina, & McHugh, 2017).
18. Among several reasons for this is that RFT assumes an intractable link between derived stimulus relations and verbal behavior, in which case any problem involving derived stimulus relations is also a problem involving verbal behavior. Verbal interventions also are compatible with the brief consultations through which clinicians often serve clients.
19. ACT theorists also assume that, through transformation of function, private events can acquire aversive properties and become the motivating operation for avoidance behaviors that reduce the private aversive stimulation but interfere with well-being behaviors. A complete explanation is beyond the scope of the present chapter.
20. This example may be easiest to appreciate by recalling from Chapter 19 that the baseline relations of stimulus classes are conditional discriminations, which means learning not just what a stimulus "is" in relational context but also what it "is not." Thus, opposition relations always are inherent in the acquisition of stimulus relations, and through a generalized opposition repertoire we are always, by experiencing pleasure, put at risk of experiencing pain when contextual cues evoke "opposite."
21. We thank Ciara McEnteggart for suggesting this example.
22. This is, perhaps, functionally similar to response interruption/redirection procedures that use "distraction" to prevent imminent stereotypical and repetitive public behaviors (e.g., Ahrens, Lerman, Kodak, Worsdell, & Keegan, 2011). Other interpretations are possible.
23. The primary literature on ACT, which any interested reader will need to consult in order to become properly conversant with ACT interventions, employs its own complex vocabulary that we mostly avoided here.

CHAPTER 21

Imitation, Modeling, and Observational Learning

LEARNING OBJECTIVES

- Define and provide examples of echoics and imitation.
- Discriminate between examples and nonexamples of imitation.
- Describe the imitation process; state what a trainer should do during the training process, based on learner performance.

Consider a trio consisting of two guitarists and a singer-songwriter performing in a contemporary rock band. During their opening set, the lead guitarist improvises a short riff. His accompanying guitarist, and band's newest and youngest member, hears the riff and then immediately reproduces it note-for-note.

Backstage between sets, the singer-songwriter says to the lead guitarist, "That was an impressive riff you dropped into our first number, but the style doesn't fit our sound. Don't play that type of riff again." The accompanying guitarist overhears the singer-songwriter's comment. During the group's second set, the lead guitarist bangs out several riffs in the same style that he played in the opening number. The accompanying guitarist does not imitate these riffs.

Imitation, modeling, and observational learning are responsible for a great deal of learning, much of it occurring without the intentional intervention by others. The two guitarists anecdote illustrates all three processes. The accompanying guitarist imitated a riff modeled by the lead guitarist. To be most adaptive, observational learning requires more than attending to and duplicating the behavior of others. Observational learners who discriminate the consequences produced by another person's behavior can better determine which behaviors to imitate. After observing the singer's disapproval of the lead guitarist's improvised riff, the accompanying guitarist stopped playing similar riffs.

The first section of this chapter defines imitation and outlines an imitation training protocol for learners who do not display imitative behavior. The chapter's second and final sections describe how applied behavior analysts use modeling as a behavior change strategy and teach children how to be skilled observational learners.

IMITATION

An imitative repertoire promotes the quick acquisition of behaviors such as a young child's development of social and communication skills. Without an imitative repertoire, a person has little chance for the agile acquisition of behaviors. Imitation has received considerable experimental and theoretical attention spanning several decades (e.g., Baer & Sherman, 1964; Carr & Kologinsky, 1983; Garcia & Batista-Wallace, 1977; Garfinkle & Schwartz, 2002; Wolery & Schuster, 1997). The experimental literature shows that individuals can acquire and maintain imitative behavior in the same fashion as they acquire and maintain other operant behaviors. That is, (a) reinforcement increases the occurrence of imitation; (b) when some imitative behaviors receive reinforcement, other imitative behaviors occur without specific training and reinforcement; and (c) some children who do not imitate can be taught to do so.

Imitation Defined

A technical definition of **imitation** entails four criteria: (a) The imitative behavior is occasioned by another person's model of the behavior (or symbolic representation thereof), (b) the imitative behavior has formal similarity with the model, (c) the imitative behavior follows the modeled behavior closely in time, and (d) the model is the primary controlling variable for the imitative behavior.

Models

A **model** is an antecedent stimulus with topographical similarity to the behavior the analyst wants to be imitated. A model can be a person who actually demonstrates a behavior (our lead guitarist playing a riff), or a model can be symbolic (i.e., a picture, video) (Bandura, 1977).

Planned models are prearranged demonstrations of behavior that show the learner exactly what to do, be it acquiring new skills or refining the topography of certain elements of existing skills.

Planned models can be live performances, as when a yoga instructor demonstrates how to achieve the adho mukha svanasana pose (a stretch in which one stands on the hands and feet; Downs, Miltenberger, Biedronski, & Witherspoon, 2015). Videos of a person performing the target behavior often serve as planned models. LeBlanc and colleagues (2003), for example, used video models to teach perspective-taking skills to three children with autism. The children learned to imitate a video model touching or pointing to objects, such as a bowl or a box. Also, the children imitated vocal-verbal behaviors, such as saying "*under the bowl*, or saying "*one.*"

Unplanned models, for better or worse, occasion imitation in everyday social environments (e.g., school, work, play, community). A tourist who has never ridden the subway system in a foreign country imitates the ticket-scanning behavior of the rider ahead of her. The turnstile opens and she enters the platform. While visiting a prospective day-care center with his parents, a child sees a series of photos of children cleaning up litter from their playground. On the way out the door, the child picks up a discarded plastic bottle and places it in the recycle bin. An undesirable outcome of imitating an unplanned model might result if an international tourist imitates a local resident jaywalking across a dangerous intersection.

Formal Similarity

Formal similarity occurs when the model's and the imitator's behavior physically resemble each other and are in the same mode (i.e., they look, or sound, alike) (Michael, 2004). For example, when a student observes a teacher fingerspell the word *house* (i.e., the model), and the student then duplicates the fingerspelling correctly (i.e., the imitation), the student's fingerspelling has formal similarity with the model. When a baby sitting in a high chair taps the food tray with his hand after seeing his mother tap the tray with her hand, the baby's tap has formal similarity to the mother's tap.

Temporal Relationship

A close temporal relationship between the model's performance and the learner's behavior is an important feature of imitation. We believe most instances of "delayed imitation" are better understood by referencing other behavioral principles and processes. The longer the latency between a previously observed model and the occurrence of behavior that duplicates the model, the more likely it is that variables other than the model are controlling the response.

Let's consider two examples that at first glance would appear to be imitation. When exiting the subway station at her destination, our international tourist emits the same turnstile operating behavior she imitated from another rider's model to gain entry to the train platform at the start of her ride. When asked by his parent, "What did you learn at school today?" a student fingerspells *h-o-u-s-e* exactly as he imitated a deaf classmate's fingerspelling the word earlier in the day.

When a response with the same topography as a previous imitation occurs in the absence of the model (e.g., another rider operating the turnstile), that response is *not* imitative behavior. The relation between a discriminative stimulus (e.g., the turnstile) that signals the availability of reinforcement because of prior experience with ticket-scans that opened a turnstile—or between a motivating operation (e.g., the parent's question) that increases the current frequency of behavior that has produced relevant reinforcement—is functionally different from the relation between an antecedent model and an imitative behavior evoked by the model. Therefore, by definition, delayed behaviors replicating the topography of a prior imitative model, are not imitative.

Model as Primary Controlling Variable

Imitation is typically viewed as *doing the same*. Although the formal similarity of *doing the same* is a necessary condition for imitation, it is not sufficient. Formal similarity can exist without the model functionally controlling the similar behavior. The most important property defining imitation is the controlling relation between a model's demonstration and a learner's performance of topographically similar behavior.

Baer, Peterson, and Sherman (1967) state: "Any behavior may be considered imitative if it temporally follows behavior demonstrated by someone else, called a model, and *if its topography is functionally controlled by the topography of the model's behavior*." (p. 405, emphasis added)

A controlling relation between the model's behavior and the imitator's behavior is inferred when a model evokes a similar behavior without a prior history of reinforcement. An imitative behavior is a *new* behavior that follows a *novel* antecedent event (i.e., the model). After the model evokes an imitative behavior, that behavior typically contacts contingencies of reinforcement. These new contingencies of reinforcement then become the controlling variables for the repetitions of the behavior (i.e., $MO \rightarrow S^D \rightarrow R \rightarrow S^R$).

Holth (2003) explained how novel instances of imitation differ from discriminated operants in this way:

> Let us imagine that a dog is trained to sit whenever the owner sits down in a chair, and to turn around in a circle whenever the owner turns around. Does the dog imitate the owner's behavior? Almost certainly not. The dog could have as easily been taught to sit whenever the owner turns around in a circle and to turn around in a circle when the owner sits down. Thus, what may look like imitation may be nothing more than a series of directly taught *discriminated operants*. The source of control can be determined only by introducing novel exemplars. In the absence of a demonstration that the dog responds to novel performances by "doing the same," there is no evidence that a similarity to the owner's behavior is important in determining the form of the dog's response. Hence, there is no true demonstration that the dog imitates the behavior of the owner unless it also responds to new instances of the owner's behavior by "doing the same." (p. 157, emphasis in original)

Holth (2003) concluded by stating that imitation "can be inferred only through instances that do not have a direct reinforcement history" (p 157). This chapter's opening anecdote included an example we infer to be true imitation. The lead guitarist improvised a short riff. The accompanying guitarist, who had never heard or played the riff, immediately reproduced

it note-for-note. The example meets the conditions for imitation: A model was present, formal similarity was achieved, a brief latency occurred between the model and the imitation, and the accompanying guitarist had no reinforcement history with the novel riff.

We'll use our two guitarists one more time to illustrate an instance of replicating another person's behavior that may appear to be imitation but is not. The lead guitarist says, "I have an idea for our opening song. Let me play it for you. If you like it, I will teach it to you." The accompanying guitarist liked what he heard. The two musicians practice the riff together until the accompanying guitarist can play it as well as the lead. Later on stage, the lead guitarist plays the riff and the accompanying guitarist immediately reproduces it. This is *not* an imitative behavior; it is an example of a discriminated operant with a history of reinforcement.

Imitation Training

An imitative repertoire enables the quick acquisition of new, complex behaviors characteristic of many human endeavors. Typically developing children acquire many skills by imitating unplanned models. Most children imitate the behavior of others without undue use of prompts or contrived reinforcement arrangements by parents, teachers, or other caregivers. Melzoff and Moore (1977) reported that infants as young as 12 to 21 days old can imitate adult hand and facial models.

For reasons that are not entirely clear, some infants and children with autism and developmental disabilities do not imitate. As a result, learning that might ordinarily progress quickly requires intensive stepwise and time-consuming instruction. Children without an imitative repertoire have little chance of rapidly developing new behaviors and great difficulty advancing beyond basic skills. Baer et al. (1967) state the importance of training such children to imitate:

> The development of a class of behaviors which may fairly be called "imitation" is an interesting task, partly because of its relevance to the process of socialization in general and language development in particular, and partly because of its potential value as a training technique for children who required special methods of instruction. (p. 405)

Imitation training is a systematic, research-based set of steps for teaching a non-imitative learner to imitate models of novel behaviors. Baer and colleagues' (1967) classic study, "The Development of Imitation by Reinforcing Behavioral Similarity to a Model," provided a pioneering demonstration of how to teach an imitative repertoire. They worked with three young children with severe to profound intellectual disabilities who did not exhibit imitative behavior. The children were taught to emit simple discriminated (i.e., similar to the model) responses (e.g., raise left arm) when the experimenter presented the verbal response cue "*Do this*," and then provided the model (e.g., raise left arm). Baer and colleagues selected appropriate skills for their participants (e.g., nod yes, stand up, put on hat, open the door). The experimenter initially used physical guidance to prompt the similar response, and then gradually reduced the guidance over several trials. The word "Good" and small bits of food were used as reinforcement to shape closer and closer topographical similarity to the experimenter's modeled responses.

The criterion for the establishment of an imitative response was the subject displaying a novel response demonstrated for the first time by the experiment with no shaping or prompting procedures for that trial.

The imitation training protocol developed by Baer and colleagues taught learners who did not imitate to imitate, meaning that a novel model controlled imitative behaviors in the absence of specific training and reinforcement of those behaviors. One participant imitated a novel model only after receiving imitation training for similarity on 130 different models. Results for the second participant were similar to those of the first. After learning just eight discriminated operants of imitative topography, Subject 3 imitated the ninth response on its first presentation. This participant imitated the ninth training model, a novel model without a history of training and reinforcement.

As a result of imitation training, two participants imitated the model with 100% similarity, but only after receiving in excess of 100 reinforced training trials. The importance of delivering positive reinforcement during and after imitation training cannot be overstated.

> The initial training procedure contained occasions when the extent of the developing imitative repertoire of each subject could be seen. These were occasions when behavior was demonstrated by the experimenter to the subject for the first time. Any attempt by the subject to imitate such new behavior before direct training or shaping could be attributed to the history of reinforcement for matching other behavior of the experimenter. (Baer, Peterson, & Sherman, 1967, p. 411)

Additionally, all three children continued to imitate new responses maintained over many sessions as long as at least some imitative responses were reinforced. When reinforcement was no longer contingent upon imitative behavior, both the previously reinforced imitations and the never reinforced probe imitations decreased markedly. However, all imitative behavior recovered when, with a small amount of shaping, reinforcement was again made contingent upon imitative behavior. Figure 21.1 shows these results for Subject 3.

The results of Baer and colleagues (1967) can be summarized as follows: (a) Children who did not have an imitative repertoire learned to imitate following a protocol that included shaping, cues, physical guidance, and reinforcement; (b) when target imitative behaviors produced reinforcement, the participants imitated other novel models without reinforcement; (c) the children demonstrated a *learning set* effect (Harlow, 1959), or a so-called *learning-to-learn* phenomenon (as the participants progressed through imitation training, they required fewer training trials to acquire new imitative repertoires); and (d) the children's imitative repertoire maintained over many sessions as long as reinforcement was contingent on at least some imitative responses.

The objective of imitation training is to teach learners to do what the person providing the model does regardless of the behavior modeled. If the training is effective, the learner

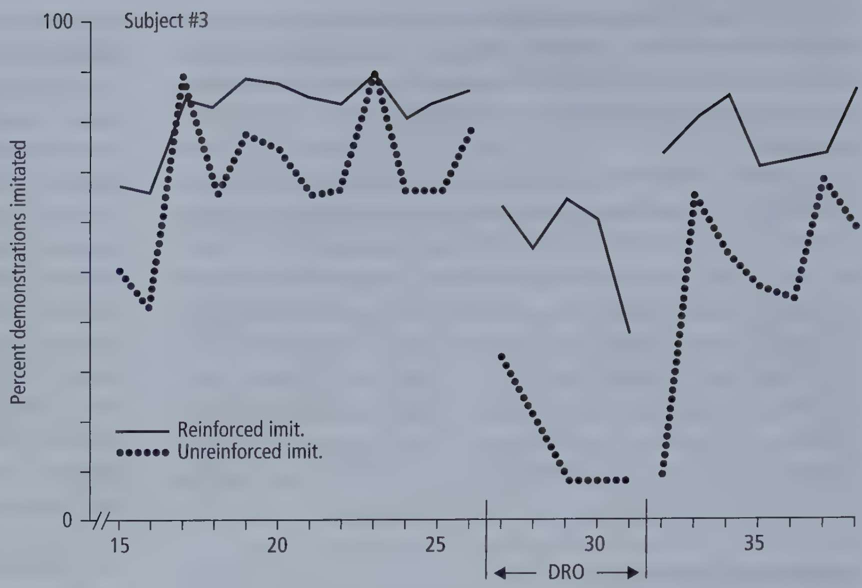


Figure 21.1 Percentage of demonstrated imitated behaviors by a child with severe/profound intellectual disabilities during reinforcement for imitation, DRO (reinforcement for any behavior other than imitation), and re-establishment of reinforcement. Solid data path shows results for eight reinforced imitations; dotted data path shows results for imitations that were never reinforced.

From "The Development of Imitation by Reinforcing Behavioral Similarity to a Model," by D. M. Baer, R. F. Peterson, and J. A. Sherman, 1967, *Journal of the Experimental Analysis of Behavior*, *21*, p. 413. Copyright 1967 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

is likely to imitate models that have not been associated with specific training, and those imitations are likely to occur in many situations and settings, frequently in the absence of planned models and reinforcement. The term **generalized imitation** is often used when a learner imitates a wide variety of unprompted, untrained, non-reinforced modeled behaviors in different settings and situations. Generalized imitation could be considered a higher order response class in which participants imitate novel models without training and that are not predictive of reinforcement.

The extent to which imitation generalizes types or classes of modeled behaviors may depend on the parameters of the response class used during training and may require programming directly. For example, Young, Krantz, McClannahan, and Poulson (1994) found that children with autism imitated novel models within the vocal, toy-play, and pantomime response types used for training, but imitation did not generalize across model types.

An Imitation Training Protocol

Building upon the experimental methods of Baer and colleagues, Striefel (1974) developed an imitation training protocol for practitioners. The components of Striefel's program are (a) assess, and teach if necessary, prerequisite skills for imitation training; (b) select models for training; (c) pretest; (d) sequence models for training; and (e) conduct imitation training.

Assess, and Teach if Necessary, Prerequisite Skills for Imitation Training

Learners cannot imitate if they do not attend to the presentation of the model. Therefore, attending to the model is a prerequisite for imitation training. Striefel (1974) defined attending as staying seated during instruction, keeping one's hands in one's lap, looking at the trainer whenever one's name is called, and looking at objects identified by the trainer. Also, practitioners often need to decrease problem behaviors that interfere with training (e.g., aggression, screaming, odd hand movements).

Suggested procedures for assessing attending skills include the following:

1. *Staying seated.* Seat the learner and record the duration of time the learner remains seated.
2. *Looking at the teacher.* Say the learner's name in a commanding voice and record whether the student makes eye contact.
3. *Keeping hands in lap.* Prompt the student to put his hands in his lap and record the duration of time the student's hands remain in that position.
4. *Looking at objects.* Place several objects on a table and say, "Look at this." Immediately following the command, move a finger from in front of the learner's eye to one of the objects and record whether the student looked at the object.

Teachers often assess attending skills for a minimum of three sessions. The teacher can begin imitation training if the assessment data show adequate attending skills. The teacher will need to teach these skills before beginning imitation training if attending skills need to be developed.

Select Models for Imitation Training

Practitioners may need to select and use about 25 behaviors as models during initial imitation training. Including gross motor (e.g., raising a hand) and fine motor movements (e.g., manual sign language) as models provides learners with opportunities to develop more refined differentiations with their imitative skills.

Practitioners usually use one model at a time during the initial training trials, rather than a sequence of movements. Practitioners may choose to use more complex models such as sequences of behavior after the learner can imitate one model per occasion successfully. Also, the initial training usually includes models of (a) the movement of body parts (e.g., touching nose, hopping on one foot, bringing hand to mouth) and (b) the manipulation of physical objects (e.g., passing a basketball, picking up a glass, zipping a coat).

Pretest

The learner's responses to the selected models should be pretested. The pretest may show that the learner will imitate some models without training. The pretesting procedures advocated by Striefel (1974) are as follows:

1. Prepare the learner's attending behaviors for the pretest (e.g., seated, hands in lap) and often assume the same ready position as the learner.
2. If you are using an object model, place one object in front of the learner and one object in front of yourself.
3. Say the learner's name to start the pretest, and when the learner makes eye contact, say, "Do this" (i.e., child's name, pause, "Do this").
4. Present the model. For example, if the selected behavior is to pick up a ball, pick up the ball yourself and hold it for a few seconds.
5. Immediately praise each response that has formal similarity to the model, and deliver the reinforcer (e.g., a hug, an edible) as quickly as possible.
6. Record the learner's response as correct or incorrect (or no response), or as an approximation of the model (e.g., touches the ball, but does not pick it up).
7. Continue pretesting with the remaining models.

Practitioners can use the pretest procedure with all motor and vocal verbal models (e.g., name, pause, "Do this," "Say ball"). Striefel recommended pretesting for several sessions until all models have been pretested at least 3 times. If the learner correctly responds during pretesting to a selected model at a set criterion level (e.g., three of three correct), then the practitioner should advance to other models. If the learner does not meet the criterion, the practitioner should select that model for imitation training.

Sequence Selected Models for Training

Practitioners use the pretest results to arrange the presentation sequence for the selected models, arranging the sequence from the easiest to the most difficult models to imitate. The first models selected for imitation training are those that the learner imitated correctly on some, but not all, of the pretest trials. The models that the learner responded to incorrectly, but that approximated the model, are selected next. Finally, the models that the learner failed to perform, or performed incorrectly, are the last to be selected for training.

Conduct Imitation Training Sessions

Striefel (1974) suggested using four conditions for imitation training sessions: preassessment, training, postassessment, and probing for imitative behaviors. The procedures used in imitation training are the same as those in the pretest, with the exception of when and how often the practitioner presents the selected models.

Preassessment. The preassessment is a short pretest given before each training session. Practitioners use the first three models currently selected for training for the preassessment. These three models are presented 3 times each in random order during the preassessment. If the learner's behavior has similarity to the model on all three presentations, that model is removed from the training. The preassessment procedure allows practitioners to evaluate the learner's current performance on the models selected for training that session, and to determine the learner's progress in learning to respond to the model.

Training. During training, practitioners use repeated presentations of one of the three models used in the preassessment. The model selected first for training is the one most often responded to during the preassessment (i.e., the behavior was similar to the model on some, but not all, of the preassessment presentations). If, however, the learner made approximations only, the behavior with the closest similarity to the model is selected first for training. Training continues until the learner responds to the model correctly on five consecutive trials.

Imitation training will likely include physical guidance to prompt the response if the learner fails to respond. For example, the practitioner may physically guide the learner's behavior through the complete response. Physical guidance allows the learner to experience the response and the reinforcer for that specific movement. After physically assisting the complete response, the practitioner will gradually withdraw the physical guidance, letting go of the student's body just before the entire movement is completed, and then continue to fade the physical guidance by withdrawing the physical support earlier on each subsequent trial. Eventually, the learner may complete the movement without assistance. When the learner responds to the model without prompting for five consecutive trials, that model is included in the postassessment.

Postassessment. During the postassessment, the practitioner presents, 3 times each, five previously learned models and five models that are still included in imitation training. The practitioner will remove a most recently learned behavior from imitation training following three consecutive postassessments in which the learner responds correctly without physical guidance to the model 14 out of the 15 opportunities. Physical guidance, however, is appropriate to use during the postassessment. If the learner does not reach this criterion (14 out of 15 postassessment opportunities), Striefel (1974) recommended continuing imitation training with that model. The postassessment procedure allows the practitioner to evaluate how well a learner performs the previously and most recently learned behaviors.

Probe for Imitation with Novel Models. Practitioners will use approximately five nontrained, novel models to probe for occurrences of imitation at the end of each imitation training session, or they will intermix the probes with the training trials. The probe procedure uses the same procedures as the preassessment activities, but without using the antecedent verbal response prompt (i.e., child's name, pause, "Do this") or other forms of response prompts (e.g., physical guidance). Probing for nontrained imitations provides data on the learner's progress in

developing an imitation repertoire—in other words, learning to do what the model does.

Guidelines for Imitation Training

Assuming that the learner has the prerequisite skills for imitation training (e.g., attending skills, existing imitative repertoire), as was the case in the Garcia et al. study, and a decision has been made as to using live versus symbolic models, the following guidelines maximize imitation training.

Keep Sessions Active and Brief

Most practitioners use short training sessions during imitation training, typically 10 to 15 minutes, but often schedule more than one session per day. Two or three 5-min sessions may be more effective than one long session. To maintain quick and active training, allow no more than a few seconds between model presentation and expected imitative response.

Reinforce Both Prompted and Imitative Responses

In the early stages of imitation training, reinforce each occurrence of prompted responses and all occurrences of imitation. If the learner's participation requires a reinforcer other than praise, use a consequence that can be presented immediately, in small amounts, and that the learner can quickly consume (e.g., cereal bits, 5 seconds of music).[1] Reinforce only matching responses or imitative behaviors that immediately follow the model (e.g., within 3 to 5 seconds). Learners who consistently emit correct matching responses, but not immediately following the model, should be reinforced for shorter response latencies (e.g., a decreasing contingency from 7 seconds to 5 seconds to 3 seconds, and so forth).

Pair Verbal Praise and Attention with Tangible Reinforcers

Often the learner's behaviors during imitation training require the use of tangible consequences such as edibles or liquids. Even so, as training progresses, it is hoped that social attention and verbal praise will maintain the matching responses or imitative behaviors. To achieve this goal, the practitioner should pair the delivery of other consequences with social and verbal praise. Social (e.g., patting the student's arm affectionately) or descriptive verbal praise should immediately follow each correct response or approximation, simultaneously with the other consequence. A learner's willingness to participate in imitation training may increase when the practitioner schedules a favorite activity to follow each session.

If Progress Stalls, Back Up and Move Ahead Slowly

There may be identifiable reasons for a worsening performance, such as reinforcer satiation or context distractions; or perhaps the practitioner presented models that were too advanced for the learner. Whether or not a clear reason exists for the worsening performance, the practitioner should immediately return to an earlier level of successful performance. Once progress resumes, training can slowly advance.

Fade Out Verbal Response Prompts and Physical Guidance

Parents and caregivers of young children almost always teach imitation skills using verbal response prompts and physical guidance. For example, the caregiver (a) tells the child to *wave bye-bye*, the caregiver models waving, and then physically guides the child's wave; (b) asks the child, "*What does the cow say?*" The caregiver presents a model, "*The cow says moo*," and then tells the child to say "*moo*"; and (c) provides praise and attention if the child says "*moo*." This natural instructional process is the same as that advocated in this chapter for teaching imitative skills to learners who do not imitate: that is, a verbal response prompt, "*Do this*"; the presentation of a model; and physical guidance when needed. However, verbal response prompts and physical guidance are not functional for producing imitation in everyday environments. Imitation training is not complete until all response prompts have been withdrawn. Children need to learn to *do what the model does* without the supports of response prompting. Therefore, to promote the effective use of imitation, fade out the response prompts used during the acquisition of trained matching responses.

Make Data-based Decisions to Terminate Imitation Training

Decisions to stop imitation training depend on the learner's behavior and program goals. For example, the practitioner could apply the criterion to terminate motor imitation training when the student imitates the first presentations of five consecutive novel models across three sessions or imitates several untrained behavior chains (e.g., washing hands, brushing teeth, fingerspelling).

MODELING

Modeling is a behavior change strategy in which learners acquire new skills by imitating demonstrations of the skills by live or symbolic models. The model shows, demonstrates, or conveys exactly the behavior the learner is expected to perform. Models can be live demonstrations or symbolic representations of the desired behavior (Bandura, 1977). A father showing his teenage son how to shave by holding and moving a razor gently across his face, and a professional golf coach demonstrating how to eliminate a hitch in a putting stroke by realigning his head and hand positions, are examples of live models. Symbolic models would be a picture icon, a photo, a video, an audio, or a mixed media presentation that illustrates or depicts the desired behavior. For example, Delli Bovi, Vladescu, DeBar, Carroll, and Sarokoff (2017) used a video model combined with voiceover instructions to teach staff how to conduct a multistep stimulus assessment procedure.

Video Modeling

Bellini and Akullian (2007) define **video modeling** as "a technique that involves demonstration of desired behaviors through video representation of the behavior. A video modeling intervention typically involves an individual watching a video

demonstration and then imitating the behavior of the model. . . . **Video self-modeling** (VSM) is a specific application of video modeling that allows the individual to imitate targeted behaviors by observing her or himself successfully performing a behavior" (p. 264).

Numerous studies have shown the positive effects of video modeling and video self-modeling across a range of ages and skills, for children and adults with various disabilities (Bellini & Akullian, 2007; Qi, Barton, Collier, & Lin, 2017; see Table 21.1). Video modeling has also been used in training applications to teach complex, multistep procedures to staff on how to conduct assessment or intervention protocols (Deliperi, Vladescu, Reeve, Reeve, & DeBar, 2015). Finally, video or self-modeling can be fashioned using standard videotape, digitally through smart pad or tablet applications, through animations, and from various camera angles and backgrounds, including so-called *point-of-view perspectives*. The fact that video modeling (VM) or VSM can be combined with instructions, feedback, and rehearsal options makes it more practical and desirable for developing new behaviors. Regardless of medium or mode, the evidence shows that video and self-modeling approaches produce imitation outcomes and, in some cases, generalized imitation outcomes.

Aldi et al. (2016) provide an example of using modeling and imitation training. They used point-of-view video models to teach daily living skills (DLS) to two 18-year-old young men with autism spectrum disorder (ASD). The two young men received training in each of their residences on three daily living skills. Participant 1 received training on cooking, setting the table, and folding jeans. Participant 2's tasks included setting the table, cleaning the bathroom sink/counter, and cleaning the mirror.

The authors used Apple iPad iOS software to create the point-of-view video models by filming family members' arms and hands doing the assigned tasks. The filming included a verbal description of each step of the tasks. Depending on the DSL to be modeled, the length of the video training tape varied, usually less than 6 minutes. Further, the number of steps that were modeled varied from 9 to 23, depending on the task (e.g., folding jeans required 9 steps, whereas cooking required 23 steps). Intervention consisted of participants watching a video related to the skill set being modeled. Subsequent to viewing the video, the participant was asked to complete the task (e.g., fold clothes). If the participant made an error, the video segment for that task was repeated as many times as necessary to achieve mastery. In some instances, gestural prompts were provided as part of an error correction procedure. Correct responses were praised.

Figure 21.2 shows that both participants mastered their respective skill sets. During a 1-month follow-up, skills maintained well above baseline levels, but were lower than the established treatment mastery levels, suggesting that additional programming for maintenance would be necessary.

Applied behavior analysts have researched the utility of a variety of models and behaviors to ascertain the effectiveness of modeling and imitation training. Overall, these advances stem from the pioneering work of Baer et al. (1967) and Bandura (1977). Table 21.1 provides a sampling of models, participants, and participants' behaviors addressed in training across daily living, safety, and academic skills.

Guidelines for Effective Modeling

It has long been established in the experimental literature that a number of factors characterize effective models (Noell, Call, & Ardoin, 2011; Sulzer-Azaroff & Mayer, 1977). The more of these characteristics that the model shares with the imitator, the more likely the modeling–imitation relationship will be established and maintained. The following sections outline several of these factors.

Model's Similarity with Learner

The similarity between the individual providing the model and the learner can influence the likelihood that imitative behavior will occur (Bandura, 1969). Almost any variable can relate to similarity: for instance, past experience, age, sex, or physical appearance. If imitation is used to teach a teenager how to reduce drug use, a person who has successfully "gone straight" might serve as an effective model, more so than someone who has not.

Prestige

A model with prestige or social status can increase the likelihood of imitative behaviors. These individuals are often leaders. For instance, high school students may imitate the dress or hair style of their favorite rock stars. Young business executives may imitate the lifestyles of their firm's senior executives. Consumers tend to buy products that are "pitched" by movie stars, celebrities, or other high-status individuals. The influence of a prestigious model is enhanced when the model and the imitator have other similar characteristics (e.g., they share music, art, or athletic interests).

Emphasize the Critical Aspect of Behavior

Emphasizing the critical aspects of the modeled stimulus can increase the likelihood of imitation. For instance, a teacher giving instruction on sight words might hold up a card with the word *blue* written on it and tell the students that the word is *blue*. The word *blue* can be emphasized by inserting a short pause between her direction, "Say . . . blue." The word can also be emphasized by destressing the intensity of the word *say* and increasing her voice volume with the word *blue*, "Say . . . BLUE." According to Risley and Reynolds (1970), emphasized stimuli increase the probability of an imitative response and boost the likelihood of imitation as a function of the emphasized model.

Instructions

Imitation can be enhanced when instructions are combined with a model's presentation. Instructions—whether spoken or written words—set the occasion for the occurrence of a specific behavior. Instructions can serve as an orienting and prompting function that increases the likelihood of imitation. A number of researchers have combined instructions, including voiceover instructions, as part of a package program to enhance the effects of video modeling (Deliperi et al., 2015).

TABLE 21.1 A Sampling of Modeling Studies Across Social, Daily Living, Safety, Academic, and Vocational Skills

Model/Format	Behavior	Learners
	Social Skills	
Self-modeling/Video (Sherer et al., 2001)	Conversation skills	Children with autism
Typically developing peer/Point-of-view video (Nikopoulos & Keenan, 2004)	Vocal or gestural initiations	Children with autism
	Daily Living Skills	
Point-of-view video modeling (Aldi et al., 2016)	Cooking, setting the table, folding jeans, cleaning bath sink/counter and mirror	Young adults with autism
Point-of-view video modeling/ Step-by-step video prompting (Cannella-Malone et al., 2011)	Starting a load of laundry and hand-washing a plate, spoon, and cup	Students with severe intellectual disabilities
Point-of-view video modeling (Shipley-Benamou, Lutzker, & Taubman 2002)	Making orange juice, preparing a letter to be mailed, setting the table, cleaning a fish bowl, feeding pet cat	Children with autism
Point-of-view step-by-step computer-presented video prompting (Sigafoos et al., 2005)	Making microwave popcorn	Adults with developmental disabilities
Live therapist modeling (Fu et al., 2015)	Food consumption	Two boys with autism and food selectivity
	Safety Skills	
Live instructor model with practice and feedback (Christensen, Lignugaris-Kraft, & Fiechtl, 1996)	First-aid skills: treating cuts	Preschoolers with disabilities
Experimenter live (Garcia, Dukes, Brady, Scott, & Wilson, 2016)	Fire safety skills	Three 4- to 5-year-old children with autism
Live and video modeling with rehearsal, praise and corrective feedback (Gunby, Carr, & LeBlanc, 2010)	Abduction-prevention responses (e.g., saying "no" when presented with a lure by a stranger, immediately running to a safe area)	Children with autism
Adult trainer/Live (Miltenberger et al., 2005)	Gun safety skills	Typically developing children ages 4 and 5 years
	Academic Skills	
Live peer vocal model (Taylor, DeQuinzio, & Stine, 2012)	Reading sight words	Children with autism
Picture prompts (Vedora & Barry, 2016)	Receptive labeling	Teenagers with autism
Video modeling (LeBlanc et al., 2003)	Perspective-taking skills	Children with autism
Step-by-step video peer model with handheld computers (Cihak, 2009)	Geometry skills	Secondary students with learning disabilities
	Vocational Skills	
Video model of scripted and naturalistic settings (Allen, Wallace, Renes, Bowen, & Burke, 2010)	Wear "Rocky the Raccoon" WalkAround costume and greet customers at retail warehouse	Adolescents and young adults with autism
Video models of peers, coworkers, job coaches performing tasks with written or voice-over instructions iPods (Kellems & Morningstar, 2012)	Vacuuming lobby, cleaning sidewalk, taking inventory of items to restock, recycling cardboard, etc.	Young adults with autism

Context

The model's presentation is more likely to be imitated if it occurs in a real situation or in a facsimile manner to the real situation. Further, these situations should be expanded to account for variations in circumstances and to promote generalized imitation.

Rehearsal and Feedback

Providing the learner with multiple opportunities to respond, combined with feedback on key elements of the model's presentation, enhances the effectiveness of modeling (Poche, Yoder, & Miltenberger, 1988). Rehearsal provides the analyst with the chance to (1) determine if the learner has acquired the imitative

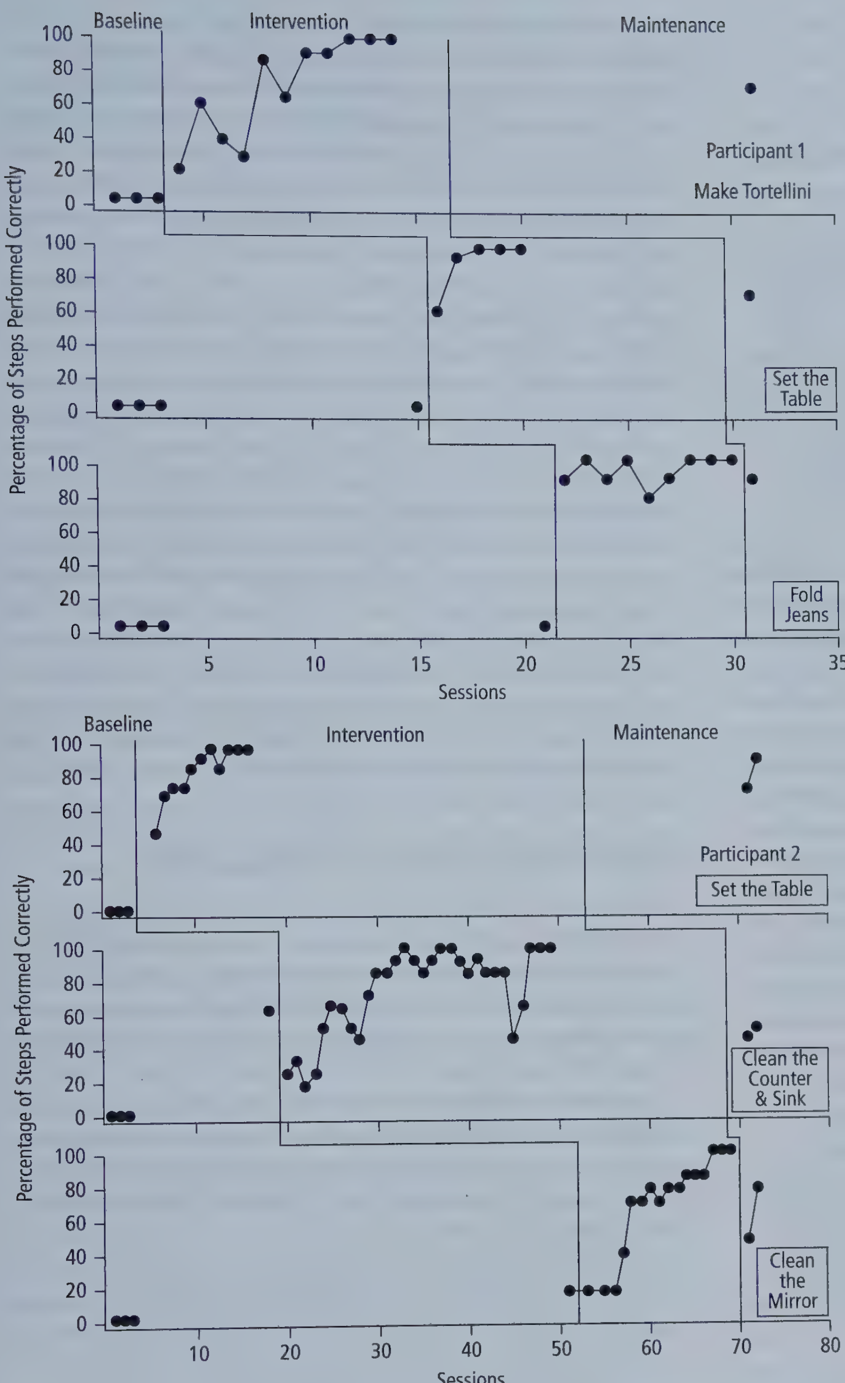


Figure 21.2 Percentage of correct daily living skills steps performed by two young men with autism.

From "Examining the Effects of Video Modeling and Prompts to Teach Activities of Daily Living Skills," by C. Aldi, A. Crigler, K. Kates-McElrath, B. Long, H. Smith, K. Rehak, and L. Wilkinson, 2016, *Behavior Analysis in Practice*, 9, p. 386. Copyright Springer International Publishing. Reprinted by permission.

skill, (2) reinforce imitative behaviors, and (3) correct any errors through feedback and redirection (Garcia, Dukes, Brady, Scott, & Wilson, 2016).

Reinforcement

The more frequently reinforcement is delivered to the imitator, the more likely imitative behavior is to increase in the future. Likewise, the more frequently the imitator observes, detects, and recognizes the model being reinforced, the more effective the model becomes (Bandura, Ross, & Ross, 1963; Noell et al., 2011).

OBSERVATIONAL LEARNING

Observational learning entails detecting another person's behavior and its consequence(s) and using that information to determine whether or not to imitate the behavior (Delgado & Greer, 2009; DeQuinzio & Taylor, 2015). Observational learning typically involves imitation but does not require it (Catania, 2013). For example, after seeing another person walking ahead of her slip on "black ice," a pedestrian may avoid that part of the sidewalk by crossing the street or stepping around it.

Does observational learning require sight? Box 21.1 explores this important question.

The objective of observational learning is that the participant "learns from indirect contact with consequences experienced by others" (Delgado & Greer, 2009, p. 408). MacDonald and Ahearn (2015) provide a classroom-related example of observational learning (OL):

> An example of OL would be when a student does not correctly point to a written word in the presence of a picture associated with the written word. The student observes another student receive instruction by a teacher followed by consequences (in the form of positive or corrective feedback). Then, because of observing this performance and the consequences that follow, the student correctly points to the word associated with the picture in the absence of direct instruction. (p. 801)

Delgado and Greer (2009) taught children with autism to discriminate the correct and incorrect sight word reading responses of confederate peers by selecting a green block if the teacher's feedback indicated the confederate peer's response was correct and selecting a red block if the teacher's feedback indicated an incorrect response. Only under conditions in which monitoring correct and incorrect feedback took place did desirable observational learning occur. The authors concluded that participants learned new sight words as a function of monitoring the correct and incorrect responses of their confederates.

It is critical that observational learners discriminate feedback received by the model. Recall that is exactly what happened in our anecdote with the guitarists. The accompanying guitarist discriminated the singer-songwriter's displeasure with the lead guitarist's improvised riff and quit playing such riffs. According to DeQuinzio and Taylor (2015), "learning this discrimination is a fundamental precondition for observational learning to occur" (p. 41).

Delgado and Greer (2009) suggest that the acquisition of observational learning skills may be a behavioral cusp:

> OL appears to be a behavioral developmental cusp. . . . a developmental cusp occurs when a child can contact contingencies that he or she could not before. . . . and it shows not only a cusp but a capability, in that the children can now learn in ways they could not before. (p. 430)

In a review of the literature, Taylor and DeQuinzio (2012) noted that children with ASD often have deficits with many of these responses. They suggested that teaching sustained attention to peer models, generalized imitation of a peer's vocal and motor responses, and discrimination of consequences could facilitate the emergence of observational learning in the natural environment. To date, however, no method exists for assessing the presence of observational learning and the skills necessary to engage in it, although studies have specifically examined teaching OL. Researchers have tried to teach OL through peer-yoked contingencies, peer monitoring, and teaching a differential observing response. Each study focused on teaching the children to attend to discriminative stimuli, to discriminate consequences related to the OL task, or both. Taylor, DeQuinzio, and Stine (2012) assessed the effects of teaching a monitoring response on the acquisition of differentially responding to words. During the training condition, participants were taught to imitate vocally the word read by the peer and to match the vocal response to the written word on the matching board. During the exposure condition, participants observed only the peer's reading response. Test conditions occurred 10 minutes after each condition to assess acquisition of the previously observed reading words. Results, though limited, suggested that all participants acquired expressive identification of words with and without monitoring.

Teaching Observational Learning Skills

As indicated earlier, many children with autism and developmental disabilities do not imitate, and observational learning is nonexistent. Children without observational learning repertoires "can learn new operants only via direct instruction, a process that is time-consuming and expensive" (Delgado & Greer, 2009, p. 431).

Three requisite skills have been identified as central to developing an observational learning repertoire: attending, imitating, and discriminating (DeQuinzio & Taylor, 2015; Greer, Dudek-Singer, & Gautreaux, 2006; Taylor & DeQuinzio, 2012; Taylor, DeQuinzio, & Stine, 2012).

Taylor and DeQuinzio (2012) and Taylor (2013) describe interventions for developing imitation and observational learning skills in persons without them. These interventions

BOX 21.1

Is Observational Learning for Sighted Persons Only?

Because observational learning connotes sight or vision, some readers may assume incorrectly that only sighted persons can profit from it. An individual who is blind uses observational learning when he attends to the auditory stimuli and consequences produced by the behavior of others and imitates those behaviors or not, accordingly. So, too, does an individual with dual sensory impairments (i.e., blind and deaf) when he senses and imitates tactile and kinesthetic stimulus products of another's behavior. Observational learning requires attention, imitation, and discrimination training regardless of sensory modes.

generally include the following three components: (a) develop attending skills; (b) teach imitation and program for generalized imitation; and (c) reinforce the discrimination of consequences—i.e., the response class "Do This, but Don't Do That"—and continued practice.

Observational learning requires several responses during and after the modeled action.

Attending

Attending means more than just being in proximity and oriented toward a stimulus. Delgado and Greer (2009) and Taylor and DeQuinzio (2012) indicated that attending at its fundamental level means that the participant, through peer- or confederate-mediated tactics, not only sustains attention toward the relevant stimulus but also makes an imitative and matching response that confirms he or she is observing that stimulus. For example, in Delgado and Greer's study the participant had to observe the confederate and then raise a differentially colored block indicating whether a correct or an incorrect response was made by the confederate. Likewise, Taylor et al. (2012) required the participant not only to repeat the stimulus word (imitation) but also to place a "chip" under that word on a separate matching board, signifying that the participant was attending to the stimulus word. In Taylor and colleagues' view, attending is a subset of a larger monitoring behavior repertoire that includes imitating a model and engaging in confirming responses to the model.

Imitating

Whether with discrete trial training (Smith, 2001), combining shaping, rehearsal, and physical guidance with imitation training (Baer et al., 1967), adding verbal description to imitation training (Jahr & Eldevik, 2002), or practicing immediate or delayed imitation with adults or peers (Taylor & DeQuinzio, 2012), teaching an imitation repertoire can be accomplished so that new, more complex repertoires of behavior develop. Taylor and DeQuinzio (2012) described how peers can be used to teach imitation:

> To teach imitation of peer responses, a teacher may initially have the child sit across from a peer at a table or in a play area. The teacher would then ask the peer to demonstrate actions (e.g., the peer is instructed to push a car back and forth) and would present an instruction to direct the child to imitate the peer (e.g., the teacher would say, "Do what John is doing"). If necessary, the teacher could guide the child to imitate the peer's action. When the child does, the teacher would provide praise and a preferred snack or toy, to reinforce the imitative response. To promote generalization, the teacher would have the peer model different actions each time, until the child can imitate novel actions without any prompting or reinforcement. (pp. 352, 353)

Discriminating

DeQuinzio and Taylor (2015) demonstrated an effective discrimination training procedure for students in whom this skill is lacking. Essentially, in their study, which included four students with autism, the researchers used a pool of unknown pictures as stimulus items. Across baseline, discrimination training, and 10-min posttest trials, they measured correct observed picture-labeling responses from reinforced and unreinforced trials between the teacher (presenting the trial) and a confederate adult model.

During a baseline condition, the teacher, who was seated across from an adult model and the participant, presented a picture with the direction, "What's this?" If the confederate adult model said the name of the picture correctly, the teacher provided praise (i.e., "That's right") and delivered a reinforcer. If the confederate, by prior arrangement, answered incorrectly, the teacher said, "I'm sorry, that's wrong," and a reinforcer on the table was removed from sight. Whether the adult model stated a correct or incorrect reply, the teacher never provided the correct picture label as part of a consequence. During the discriminating training phase, baseline conditions were repeated, except that the teacher, after providing a consequence to the model ("That's right)", turned immediately to the participant and said, "What's this?" If the participant answered correctly, praise was delivered. If the participant emitted an incorrect response, the teacher provided error correction in the form of a statement, "Say what she says when she gets it right." Correspondingly, after nonreinforced trials to the adult model, the picture was again presented to the participant immediately, with the statement "What's this?" If the participant said, "I don't know," or didn't say anything, that trial ended, and the next was initiated. Finally, generalization sessions were conducted approximately every three to five sessions, using different pictures than were used during baseline and discrimination training.

Figure 21.3 shows the results of the study. Each participant learned to discriminate reinforced responses of the confederate adult model and to say "I don't know" in the presence of nonreinforced responses to that model. The discrimination performance of the participants was maintained, albeit at a limited level, for a period of time after training using different pictures, but improved when additional training for generalization was instituted.

If any of the requisite skills for observational learning—attending, imitation, delayed imitation, and consequence discrimination—are lacking in a participant's repertoire, these skills can be taught separately and then combined to improve or enhance observational learning.

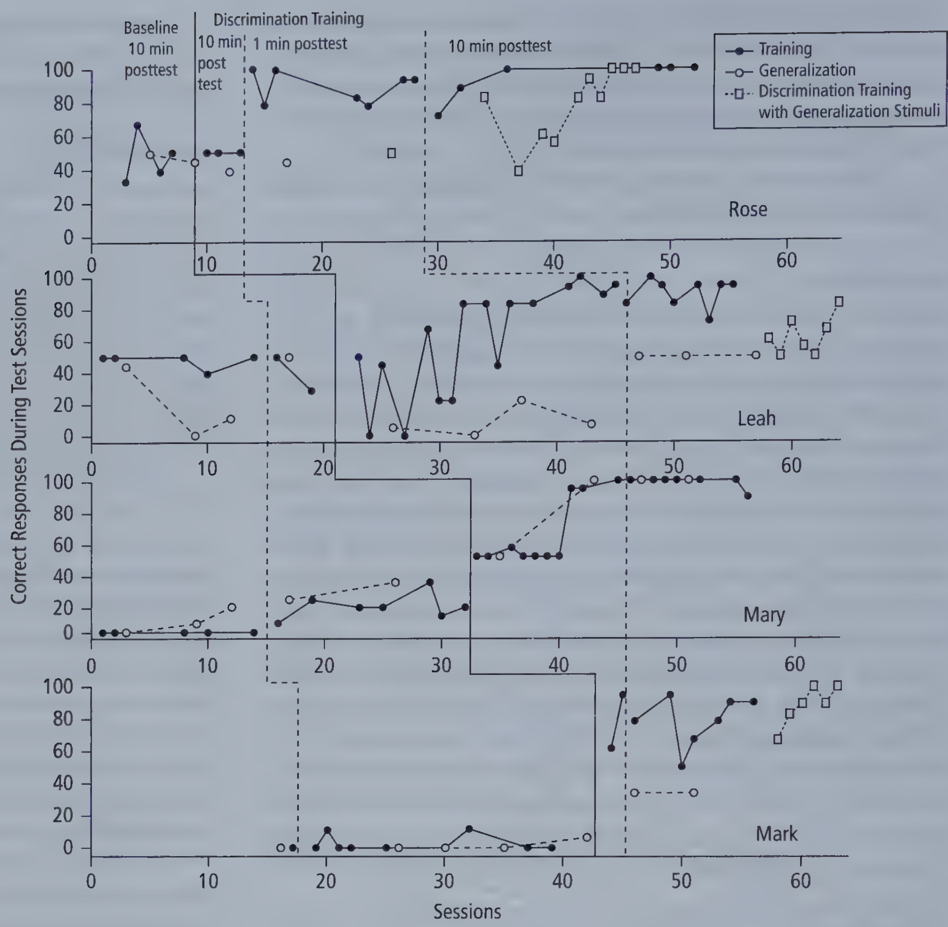


Figure 21.3 The percentage of correct responses during test sessions conducted after baseline, discrimination training, and generalization sessions for three children with autism.

From "Teaching Children with Autism to Discriminate the Reinforced and Nonreinforced Responses of Others: Implications for Observational Learning," by J. A. DeQuinzio and B. A. Taylor, 2015, *Journal of Applied Behavior Analysis, 48*, p. 46. Copyright by the Society of the Experimental Analysis of Behavior, Inc. Reprinted by permission.

SUMMARY

Imitation

1. Imitation entails four criteria: (a) The imitative behavior is occasioned by another person's model of the behavior (or symbolic representation thereof), (b) the imitative behavior has formal similarity with the model, (c), the imitative behavior follows the modeled behavior closely in time, and (d) the model is the primary controlling variable for the imitative behavior.

2. A model is a live or symbolic demonstration of the target behavior that shows the learner exactly what to do.

3. Formal similarity occurs when the imitator's behavior physically resembles the model and is in the same mode.

4. A brief latency between the model and the learner's behavior is an important feature of imitation.

5. Most instances of *delayed imitation* may be better understood with other behavioral principles and processes.

6. The most important property defining imitation is the controlling relation between a model's demonstration and a learner's performance of topographically similar behavior.

7. Imitation training is a systematic, research-based set of steps for teaching a non-imitative learner to imitate models of novel behaviors.

8. Generalized imitation refers to a learner imitating a wide variety of unprompted, untrained, non-reinforced modeled behaviors in different settings and situations.

9. Striefel's (1974) imitation training protocol for practitioners has the following components: (a) assess, and teach if necessary, prerequisite skills for imitation training; (b) select models for training; (c) pretest; (d) sequence models for training; and (e) conduct imitation training.

10. Guidelines for imitation training include the following:
    - Keep sessions active and brief.
    - Reinforce both prompted and imitative responses.
    - Pair verbal praise and attention with tangible reinforcers.
    - If progress stalls, back up and move ahead slowly.
    - Fade out verbal response prompts and physical guidance.

- Measure and record the learner's performance and review the data after each session.
- Terminate imitation training when the learner meets specified performance criteria such as: imitates the first presentations of five consecutive novel models across three sessions, imitates several untrained behavior chains (e.g., washing hands, brushing teeth, fingerspelling).

Modeling

11. Modeling is a behavior change strategy in which learners acquire new skills by imitating demonstrations of the skills by live or symbolic models.

12. Video modeling is a technique that involves the demonstration of desired behaviors through visual representations of the behavior.

13. In video self-modeling, the learners observe a video of themselves successfully performing the target behavior and imitates their self-model.

14. The effectiveness of modeling may be enhanced by:

- Models with characteristics similar to those of the learner.
- Models the learner considers to have prestige or social status.
- Models' emphasis on critical aspects of the target behavior.
- Instructions orienting and prompting the learner to attend to the model.
- Model demonstrating the target behavior in an authentic context.
- Rehearsal and feedback.
- The learner observing the model receiving reinforcement for the target behavior; the learner receiving reinforcement for imitating the target behavior.

Observational Learning

15. Observational learning entails another person's behavior and its consequence(s) and using that information to determined whether or not to imitate the behavior. Observational learning typically involves imitation but does not require it.

16. Three requisite skills are central to developing an observational learning repertoire: attending, imitating, and discriminating. If any of the skills are lacking, they can be taught separately and then combined.

KEY TERMS

generalized imitation
imitation
imitation training
modeling
observational learning
video modeling
video self-modeling

MULTIPLE-CHOICE QUESTIONS

1. With an understanding of the imitation process, applied behavior analysts can use imitation as an intervention to:

a. Eliminate old behaviors

b. Evoke new behaviors

c. Create teaching sequences

d. Reinforce appropriate behaviors

Hint: (See "Introduction to the chapter")

2. The controlling variable for an imitative behavior is:

a. An operative behavior

b. An imitative response chain

c. A model

d. A pre-task

Hint: (See "Imitation Defined")

3. Learning to purchase a soda from a vending machine by first watching someone else purchase a soda would be an example of:

a. Planned echoic stimuli

b. Unplanned echoic stimuli

c. Planned models

d. Unplanned models

Hint: (See "Imitation Defined")

4. When a child picks up a fork immediately after observing her father pick up a fork, this imitative behavior has:

a. Planned similarity

b. Unplanned similarity

c. Formal similarity

d. Informal similarity

Hint: (See "Imitation Defined")

5. Regardless of the behavior modeled, the objective of imitation training is for the learner to do what?
   a. What the model does
   b. What the trainer asked or commanded
   c. What behavior best fits the situation
   d. What behavior will receive reinforcement

   Hint: (See "Imitation Training")

6. Potential learners cannot imitate if they do not have this prerequisite skill:
   a. The behavior in their repertoire
   b. Attending to the model
   c. Decrease in problem behavior
   d. Ability to define the behavior

   Hint: (See "Imitation Training")

ESSAY-TYPE QUESTIONS

1. Four behavior environment relations functionally define imitation. What are they and why must each occur?

   Hint: (See "Imitation Defined")

2. A learner has played the game "Simon Says" many times in the past (left foot up, right foot up, left hand up, right hand up, etc.). During a game of "Simon Says", the trainer raises his right hand up without saying "Simon Says"; immediately after this model, the learner raises his hand. Is this an imitative behavior? Explain your answer.

   Hint: (See "Imitation")

3. During imitation training, if a learner fails to respond, what should a trainer do and how will this help to increase the potential for imitation learning?

   Hint: (See "Imitation Training")

NOTES

1. Before using food as a reinforcer, practitioners should determine the learner has no medical conditions that might preclude delivering food as a reinforcer (e.g., food allergies, swallowing difficulties) and should obtain permission from parents/guardians.

CHAPTER 22

Shaping

LEARNING OBJECTIVES

- Define and provide examples of shaping.
- Explain the process of shaping in terms of the behavioral contingencies used.
- Differentiate between shaping across different and within the same response topographies.
- Explain ways in which the efficiency of shaping can be increased.
- State the guidelines for implementing shaping effectively.

An Amazing Display!

In the summer of 1955 I was exploring some simple schedules of reinforcement in normal preschool children, using a variety of contingently delivered trinkets and bits of candies and cookies. The research was being conducted in a newly designed and equipped trailer laboratory at the University of Washington. In the far wall of the laboratory room as one entered was a large one-way glass, and beneath it a lever for responding (actually an O'Cedar mop handle). Above the lever was a recessed light bulb and below the lever was a small opening for delivering "goodies."

A child's responses on the lever were carried to a circuit which was programmed so that each reinforced response was accompanied by a light flash and a short buzz from the dispenser motor, indicating that a goodie was being delivered. The experimenter operated behind the wall with the one-way glass and apparatus.

One day I invited Bill Verplanck, a visiting professor in the Department of Psychology at the University of Washington, to join me in observing the performances of the children. He was impressed with what he saw: children responding in a fairly orderly fashion to fixed and variable ratio and interval schedules of reinforcement.

About halfway through the session, Verplanck suggested that we explore the shaping of operant behavior. We decided to see whether we could establish pulling the cord of the window shade on the window at the far side of the room, which was about fourteen feet away from the wall where the lever was situated. The procedure consisted of hand-shaping successive approximations to the final task, i.e., reinforcing getting off the chair, moving farther and farther away from the lever, raising the hand, grasping and then pulling on the window shade cord. Initial movements away from the lever were a bit slow in developing because each time the child was reinforced for moving away, he had to turn back or run back to get the goodie lodged in the receptacle. In a relatively short time, however, the response was well established: After snatching up a goodie, the child would run across the room, pull on the shade cord, and as soon as he heard the motor buzz, he would dash back to the receptacle and retrieve the goodie.

This was an amazing display! . . . Delivering reinforcers contingent on the topography of successive behaviors clearly showed how new operant behavior—any kind of new operant behavior—is probably established. I had just witnessed a dramatic demonstration of what we had for years been telling our students in introductory psychology, namely, that the implication of a scientific analysis of behavior is prediction and control of individual behavior.

—Adapted from "In Memoriam: Sidney W. Bijou, 1908–2009," by M. Dougher, 2009, *The Behavior Analyst*, pp. 363–364. Originally appeared in *Behavior Modification: Behavioral Approaches to Human Problems*, by W. H. Redd, A. L. Porterfield, and B. L. Andersen, 1978. New York: Random House. Used with permission.

Shaping is used in numerous situations to help learners acquire new behaviors. Wolf, Risley, and Mees (1964), for example, employed shaping in an applied setting to increase eyeglass wear for a child who was in danger of losing his eyesight if he did not wear corrective lens regularly. Athens, Vollmer, and St. Peter Pipken (2007) used shaping to improve the academic behavior of students with speech, language, and learning disabilities. Researchers and trainers have used shaping to teach functional behavior to animals (e.g., walking horses into a trailer without injury to the horses or the attendants), for appeal or utility purposes (e.g., teaching porpoises to execute show routines), for humanitarian reasons (e.g., conducting mine sweeping operations in formerly embattled combat areas), for security reasons (e.g., sniffing out explosives in airports, terminals, or high-traffic public venues), for medical assessments (e.g., detection of tuberculosis or necrotic tissue), and for assisting animals to remain calm during immunization protocols

and dental examinations and procedures (Poling et al., 2017; Pryor & Ramirez, 2014).

Depending on the complexity of a given behavior, or the gap between the learner's present level of performance and the desired target (terminal) behavior, shaping may require numerous successive approximations before the terminal behavior is achieved. Acquiring the terminal behavior with respect to time, trials, or direction is seldom predictable, immediate, or linear. It is possible that a learner—human or animal—might emit a closer approximation to the terminal behavior that a practitioner fails to detect and reinforce. In this situation, achievement of the terminal behavior will be delayed. However, if a systematic approach is used—that is, if each instance of a closer approximation to the terminal behavior is detected and reinforced while simultaneously placing prior behaviors on extinction—progress can be attained more quickly.

Although shaping a behavior can be time-consuming, it is an important tactic for teaching new behaviors (Fonger & Malott, 2018). Shaping should be in every behavior analyst's repertoire, especially in those instances where it is unlikely that novel target behaviors will be learned by following instructions, gaining incidental exposure, or providing models, physical cues, or verbal prompts.

This chapter defines shaping, differential reinforcement, and successive approximations; presents examples of how to shape behavior across and within different response topographies; and suggests ways to improve shaping efficacy. Clicker training is illustrated as a method that trainers use to shape new behaviors in animals and humans. Next, emerging applications of shaping, and how shaping can be applied to large-scale social and humanitarian endeavors, is presented. Finally, we provide guidelines for implementing shaping and conclude the chapter with "how to learn to shape behavior."

SHAPING DEFINED

Shaping consists of a three-part process whereby the analyst (a) detects a change in the learner's environment, (b) makes a discriminated judgment about whether that change is a progressively closer approximation to a terminal behavior of interest; and then (c) differentially reinforces that closer successive approximation. In short, **shaping** is defined as the differential reinforcement of successive approximations toward a terminal behavior. Detecting means that the shaper must use one or more of his or her physical senses to see, hear, or otherwise note that a change has occurred. Discriminated judgment means that the shaper must answer the following question: "Is what I just saw (or heard) a closer approximation to the desired terminal behavior or not?" **Differential reinforcement**, as applied in shaping, refers to presenting an unconditioned or conditioned reinforcer only to those emitted members of a response class that share a specified dimension or quality, while placing all other emitted response class members on extinction.[1]

In other words, the shaper, after detecting a change, must decide if that change is in the direction of an ever-evolving nearer approximation to the desired terminal behavior. If so, the analyst reinforces it immediately. If, after detecting a change, the analyst deems that the change occurred, but that it was not a closer approximation to the terminal behavior, then that previously emitted behavior is placed on extinction, meaning the analyst does not reinforce it.

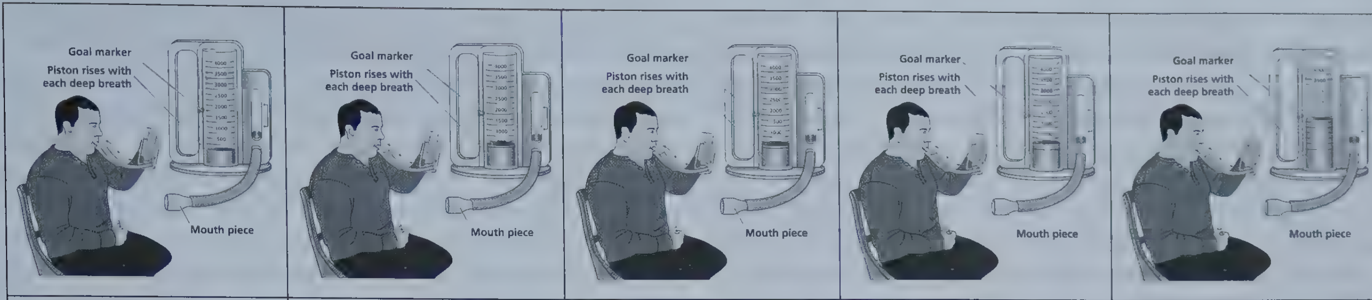
For example, assume that a respiratory therapist (RT) is charged with increasing the lung capacity of a patient recovering from pneumonia. She uses a spirometer, which measures lung capacity volume in milliliters (ml). The physician directing the treatment sets the patient's goal for 1750 ml, but the patient's present lung volume breathing behavior registers only 250 ml (i.e., a volume gap of 1500 ml). To use shaping, the RT first closely observes the patient to note the occurrence of a breathing volume change—any change. If any progressive breathing attempt exceeds 250 ml by any volume level (e.g., 250.01, 251, 255 ml) praise is delivered ("Well done, Yuri! That's higher!"). Any lower volume inhalation response level is placed on extinction.[2] Figure 22.1 shows a series of four initial trials illustrating this process as the patient works his way toward his goal.

Behavior analysts should not be deceived, however, by the apparent *step-wise* progression of improved breathing levels in Figure 22.1 (e.g., 300 ml, 325 ml, 500 ml). These levels were arbitrarily provided as a pedagogical illustration. Shaping, in its procedurally accurate application, is absolutely not a step-wise, fixed-gradient process that would require the patient to go from 250 ml to 300 ml to 350 ml and so on upward to the goal of 1750 ml. Any breathing behavior that moved the volume indicator level incrementally higher toward the physician-directed goal of 1750 ml (say, 250.1 ml to 276.5 to 339, etc.) would meet the criterion of being a successive approximation to 1750 ml and is reinforced. Any attempts lower than an already established higher level would be placed on extinction.

In *Science and Human Behavior*, Skinner (1953) presented the concept of shaping with an analogy:

> Operant conditioning shapes behavior as a sculptor shapes a lump of clay. . . . The final product seems to have a special unity or integrity of design, but we cannot find a point at which this suddenly appears. In the same sense, an operant is not something which appears full grown in the behavior of the organism. It is the result of a continuous shaping process. (p. 91)

Through careful and skillful manipulation of the original, undifferentiated lump of clay, the artisan keeps some aspects of the clay in its primary location, cuts other pieces away, and reforms and molds still other sections so that the form is slowly transfigured into the final sculpted design. Similarly, the skillful practitioner shapes novel forms of behavior from responses that initially may bear little resemblance to the final product. To restate, a practitioner using shaping differentially reinforces successive approximations toward a terminal behavior, while simultaneously placing on extinction attempts that do not progress toward the terminal behavior. The end product of shaping—*a terminal behavior*—can be claimed when the topography, rate, latency, duration, interresponse time, or magnitude of the target behavior reaches a predetermined criterion level (e.g., the physician-determined 1750-ml lung capacity level in our example). The two key procedural components of



Baseline trial: (250 ml). The patient's current breathing level.	**First attempt after baseline:** (300 ml). Reinforcement delivered. Had this attempt been any level lower than 250 ml, even fractionally, that breathing response would have been placed on extinction.	**Second attempt after baseline:** (325 ml). Reinforcement delivered. Had this attempt been any level higher than 300 ml, even fractionally, S^R would have been delivered. However, had this attempt been any lower than 300 ml, that attempt would have been placed on extinction.	**Third attempt after baseline:** (500 ml). Reinforcement delivered. Had this attempt been any level higher than 325 ml, even fractionally, S^R would have been delivered. However, had this attempt been any lower than 325 ml, that attempt would have been placed on extinction.	**Future attempts** continue in like fashion until the 1750 ml is attained. That is, breathing attempts are differentially reinforced. Higher levels are reinforced. Lower levels are placed on extinction.

Figure 22.1 Illustrates the process by which a respiratory therapist (a) detects a change in breathing volume, (b) makes a judgment about whether that change is a successive approximation in the direction of the terminal behavior (i.e., 1750 ml); and (c) differentially reinforces that closer successive approximation.

shaping—differential reinforcement and successive approximations—are presented next.

Differential Reinforcement

> *When a heavy ball is thrown beyond a certain mark, when a horizontal bar is cleared in vaulting or in jumping, when a ball is batted over the fence (and when, as a result, a record is broken or a match or game won), differential reinforcement is at work.*
>
> —B. F. Skinner (1953, p. 97)

To reiterate, differential reinforcement is a procedure in which reinforcement is delivered for responses that share a predetermined dimension or quality, and reinforcement is withheld for responses that do not demonstrate that quality (i.e., these responses are placed on extinction). Differential reinforcement has three fundamental effects: (a) Responses within a response class similar in function to those that have been reinforced in the past are more likely to reoccur; (b) responses resembling the unreinforced members are not as likely to reoccur (i.e., these responses undergo extinction); and (c) when responses are unreinforced and undergo extinction, extinction-induced variability may occur. **Extinction-induced variability** is the phenomenon of non-reinforcement of a response class whereby a temporary increase in the *changeability* of the response class topography occurs, and which may produce, by chance alone, a closer approximation to the terminal behavior. Extinction-induced variability addresses a change in the *form* of the response, whereas extinction burst addresses a change in the *rate* of the response.

When differential reinforcement is applied consistently within a response class, its triple effect results in a new response class composed primarily of responses sharing the characteristics of the previously reinforced subclass. This emergence of a new response class is called **response differentiation**. Catania (2013) provides a succinct description of how extinction-induced variability works with respect to establishing response differentiation.

> Shaping is based upon *differential reinforcement*. At successive stages, some responses are reinforced and others aren't. In addition, the criteria for differential reinforcement change as responding changes, in *successive approximations* to the response to be shaped. The property of behavior that makes shaping effective is that behavior is variable. No two responses are the same, and reinforcement of one response produces a spectrum of responses, each differing from the reinforced response along such dimensions as topography (form), force, magnitude and direction. Of these responses, some are closer to the response to be shaped than others and may be selected to be reinforced next. Reinforcing these responses produces still others, some of which may come even closer to the response to be shaped. Thus, reinforcement can be used to change the spectrum of responses until the one to be shaped occurs. (p. 114)

Successive Approximations

A **successive approximation** refers to the gradual and progressive criterion change for delivering reinforcement for

a behavior that is closer to the terminal behavior than the response it replaces. In other words, once a more incremental behavior occurs that resembles the terminal behavior more closely, reinforcement is delivered. When an initially reinforced response is established, the practitioner shifts the criterion for reinforcement to responses that are even a closer approximation of the terminal behavior. However, care must be applied in shifting the criterion. If the criterion is set too low, an excessive number of behaviors at the same behavior-performance level will be reinforced, stalling progress. If the criterion is set too high, extinction occurs and progress is thwarted (Hanley & Tiger, 2011; Noell, Call, & Ardoin, 2014). Skinner (1953) discussed the critical nature of successive approximations as follows:

> The original probability of the response in its final form is very low; in some cases, it may even be zero. In this way, we build complicated operants that would never appear in the repertoire of the organism otherwise. By reinforcing a series of successive approximations, we bring a rare response to a very high probability in a short time. This is an effective procedure because it recognizes and utilizes the continuous nature of a complex act. (p. 92)

Figure 22.2 illustrates a progression of successive approximations that might be used to shape drawing uniformly spaced horizontal lines. Inset 1 contains products of the learner's current repertoire (Skinner's "undifferentiated lump"). Inset 6 shows products of the desired terminal behavior. Shaping begins with the learner's current behavior; in this case, by detecting and then reinforcing longer scribbles (star-marked scribbles in Inset 1) and placing shorter scribbles on extinction. The resulting response differentiation produces a new response class (i.e., longer scribbles) (see Inset 2). The shaper shifts the criteria for reinforcement to more closely approximate the terminal behavior—in this example, longer scribbles that do not cross themselves. Insets 3 through 5 illustrate the new response classes that might emerge from differential reinforcement applied to each previous response class.

Shaping Different Dimensions of Performance

Behavior can be shaped along any of its measurable and malleable dimensions (see Table 22.1).

Let's focus on magnitude for a moment—the last dimension presented in Table 22.1—as it relates to teaching a child to talk within a conversational decibel range. Let us suppose a behavior analyst is working with a student who talks at such a low volume (e.g., below 45 decibels) that his teacher and peers have difficulty hearing him. Projected and planned intermediate approximations to 65 decibels (dB)—the amplitude of normal conversational speech—might be 45 dB, 55 dB, and ultimately 65 dB. With differential reinforcement, any response at or above 45 dB (45.1 dB, 47.7 dB, 50.3 dB) would be reinforced, whereas speaking responses that occurred at 45 dB or lower would be placed on extinction. With successive approximations, the criterion for reinforcement would be raised gradually and incrementally to 55 dB, and finally 65 dB. In effect, lower magnitude response levels are

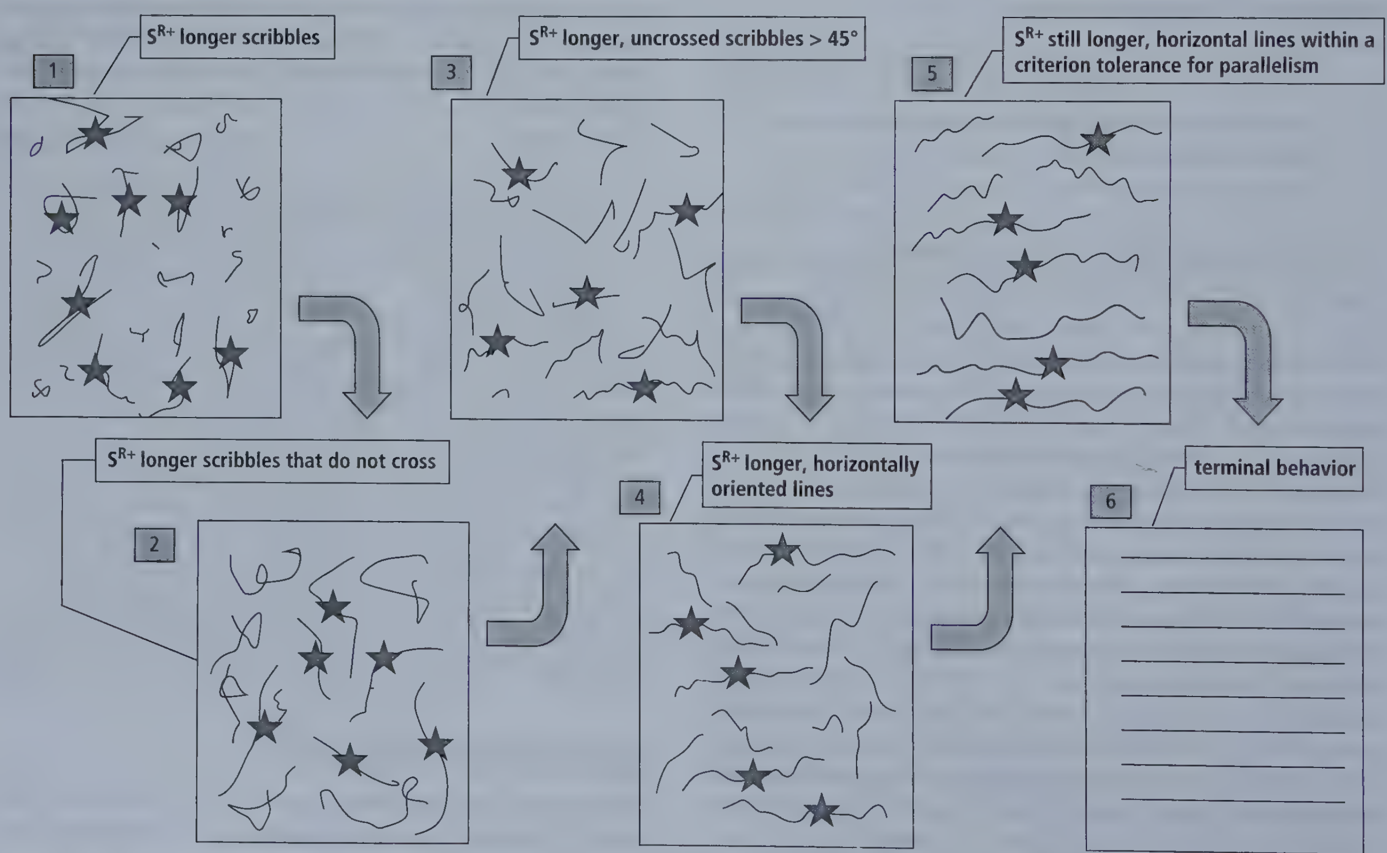


Figure 22.2 Shaping scribbles into uniformly spaced horizontal lines. Stars mark reinforced responses.

TABLE 22.1 Performance Dimensions That Can Be Shaped

Performance Dimension	Example
Topography (form of the behavior)	• Refining motor movements associated with a golf swing, throwing motion, or vaulting behavior. • Improving cursive or manuscript letter formation during handwriting exercises.
Rate (number of responses per unit of time)	• Increasing the number of problems completed per minute during a math seatwork assignment. • Increasing the number of correctly spelled and appropriately used words written per minute.
Latency (time between the onset of the antecedent stimulus and the occurrence of the behavior)	• Decreasing compliance time between a parental directive to "clean your room" and the onset of room-cleaning behavior. • Increasing the delay between the onset of an aggressive remark and retaliation by a student with severe emotional disabilities.
Duration (elapsed time from the onset of a response to its end point)	• Increasing the time a student stays on task. • Increasing the number of minutes of engaged study behavior.
Interresponse time (IRT)	• Reducing rapid eating by increasing the time (IRT) between food bites. • Reducing smoking by increasing the time (IRT) between lighting a cigarette.
Magnitude (response strength or force)	• Increasing the projected voice volume from 45 dB to 65 dB. • Increasing the required torque on a ratchet wrench from 20 foot pounds to 30 foot pounds.

not reinforced; higher magnitude responses (i.e., those that are successively above 45 dB, 55 dB, and 65 dB) are reinforced.

Fleece and colleagues (1981) used shaping to increase the voice volume of two children enrolled in a private preschool for students with physical and developmental disabilities. Baseline data were collected on voice volume in the regular classroom. Voice volume was measured on a 0- to 20-point scale, with 0 indicating that the child's voice level was inaudible, 10 indicating normal voice volume, and 20 indicating screaming. The shaping procedure consisted of having the children recite a nursery rhyme in the presence of a voice-activated relay device, whereby voice volume activated a light display. The intensity of the light corresponded to increased levels of voice volume: Higher voice volume produced a brighter light, and lower voice volume produced a dimmer light. The teacher shaped voice volume by increasing the sensitivity threshold of the relay device. Each raising of the volume level required to activate the light represented a successive approximation to the terminal volume.

Figure 22.3 shows an analysis of the children's improved performance. Also, the children's voice volume remained high after a 4-month period. Finally, according to anecdotal reports by school staff, the children's higher voice volume generalized to other settings beyond the classroom.

Applied behavior analysts can measure voice volume and apply differential reinforcement with advanced digital recording technology. Edgerton and Wine (2017), for example, used a digital tablet and app (Voice Meter Pro™) to measure the magnitude of a boy's utterances that frequently fell below conversational volume. The app displayed a thermometer-style gauge and three emoji faces (e.g., neutral, positive, and negative) that coupled visually signaled calibrated changes in voice magnitude with text prompts ("Speak up, I can't hear you!," "That's better," and "That's too loud!"). Although Edgerton and Wine did not use shaping—they implemented a package intervention of differential reinforcement, visual feedback, and corrective statements—the app they employed was adjustable and could be used in shaping voice volume.

SHAPING ACROSS AND WITHIN RESPONSE TOPOGRAPHIES

Shaping behavior across different response topographies means that selected members of a response class are differentially reinforced, whereas members of other response classes are placed on extinction. Many complex behaviors comprise component response classes, each class a necessary, but insufficient, contributor to the required performance. Speaking, for example, involves coordinated tongue and lip movements, speech sounds, one-word utterances, phrase or sentence productions, and so forth. When shaping behavior across different response topographies, the practitioner gradually increases the criterion of performance before delivering reinforcement.

Isaacs, Thomas, and Goldiamond (1960) reported a classic study showing how behaviors can be shaped across and within response topographies. They successfully shaped the vocal behavior of Andrew, a man diagnosed with catatonic schizophrenia, who had not spoken for 19 years despite numerous efforts to encourage speech production. Essentially, the shaping procedure was initiated when an astute psychologist noticed that Andrew's passive expression changed slightly when a package of chewing gum inadvertently dropped on the floor. Realizing that gum might be an effective reinforcer for building behaviors in the response class of talking, the psychologist selected speech production as the terminal behavior.

The next decision was to select an initial behavior to reinforce. Lip movement was chosen because the psychologist noted that slight lip movements had occurred in the presence of the pack of gum, and more important, lip movement was in the response repertoire of speech. As soon as lip movement was established by differential reinforcement, the psychologist waited for the next approximation of the terminal behavior. During this phase, lip movement alone was no longer reinforced; only lip movements with sound produced the reinforcement. When Andrew began making guttural sounds,

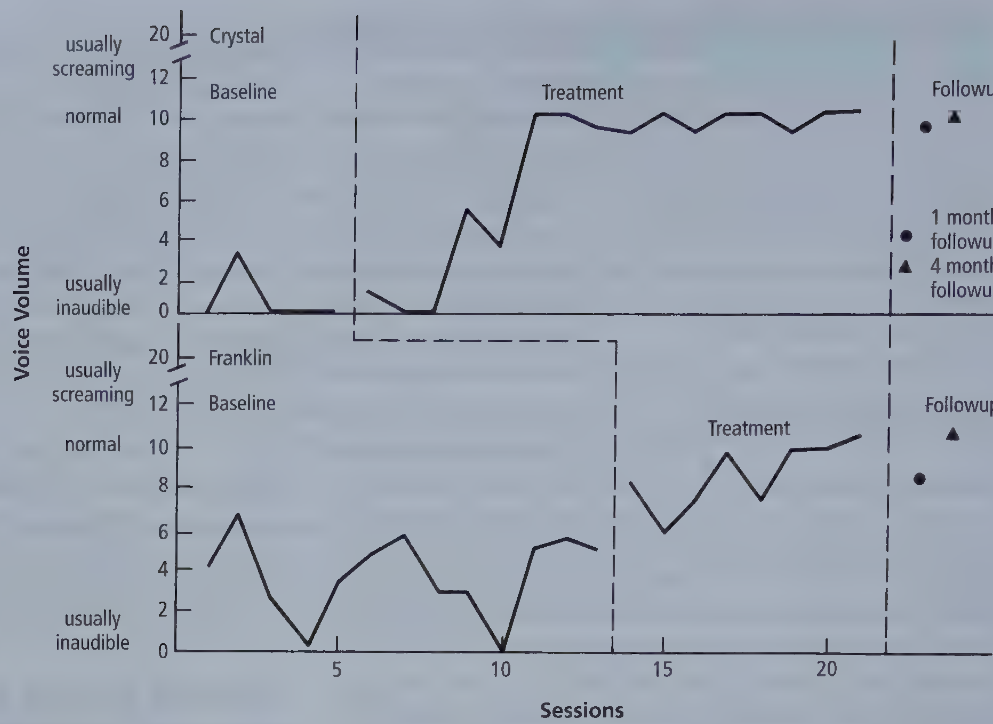


Figure 22.3 Voice-volume levels per session in the classroom generalization setting.

From "Elevation of Voice Volume in Young Developmentally Delayed Children via an Operant Shaping Procedure," by L. Fleece, A. Gross, T. O'Brien, J. Kistner, E. Rothblum, and R. Drabman, 1981, *Journal of Applied Behavioral Analysis, 14*, p. 354. Copyright by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

vocalizations were differentially reinforced. Then the guttural sound itself was shaped (differential reinforcement within a response topography) until Andrew said the word *gum*. After the 6th week of shaping, the psychologist asked Andrew to say *gum*, to which Andrew responded, "Gum, please." During that session and afterward, Andrew went on to converse with the psychologist and others at the institution about his identity and background. In this powerful demonstration of shaping, after selection of the terminal behavior and the initial starting point, each member of the response repertoire was shaped by differential reinforcement of successive approximations to the terminal behavior.

Shaping a behavior within a response topography means that the form of the behavior remains constant, but differential reinforcement is applied to another measurable dimension of the behavior (Noell et al., 2014). To illustrate, let us suppose that in a college-level physical education class, the teacher is instructing the students on water rescue. Specifically, she is teaching them how to throw a life preserver a given distance to a person struggling in the water. Since the important skill in this activity is to throw the life preserver near the person (i.e., within reach), the physical education teacher might shape accurate tossing by reinforcing successive approximations to a toss of a given distance. In other words, each toss that is near the person (e.g., within 2 meters) will be praised, whereas tosses outside that range will not be reinforced. As students become more accurate, the area is reduced so that the terminal behavior—tosses within arm's length of the person—is achieved. In this case, the magnitude of the behavior is being shaped; the form of the toss remains the same. It is important to remember that the behavior being shaped is not tossing per se. What is being shaped by differential reinforcement is the desired dimension of tossing behavior within the response class, namely, the magnitude of tossing the preserver.

Beneficial Aspects of Shaping

Shaping uses a positive (i.e., reinforcing) approach to teach new behaviors. Also, it is systematic: Reinforcement is delivered consistently upon the occurrence of any successive approximations to the terminal behavior, whereas non-approximations are placed on extinction. Punishment or other aversive procedures are typically not involved in a shaping program. Further, shaping is usually combined with other established behavior change or behavior-building procedures (e.g., verbal prompts or chaining).

Harrison and Pyles (2013) combined verbal prompts with shaping to improve tackling by high school football players. Figure 22.4 shows the results for three players during football practice and game conditions where speed was gradually shaped from walking to jogging to running.

Shaping might further be combined with a chaining procedure when a given behavior within the chain is nonexistent or underdeveloped.[3] Suppose that a behavior analyst designed a seven-step task analysis to teach a child to tie his shoes. However, the child was unable to complete Step 5 in the task analysis. Shaping could be used in isolation to teach a closer approximation to that step. Once Step 5 was learned through shaping, chaining would continue with the rest of the steps in the task analysis sequence.

Limitations of Shaping

At least five limitations of shaping can be identified. Practitioners should be aware of these limitations and be prepared to deal with them if they arise.

1. Shaping new behavior can be time-consuming because numerous approximations may be necessary before the terminal behavior is achieved.

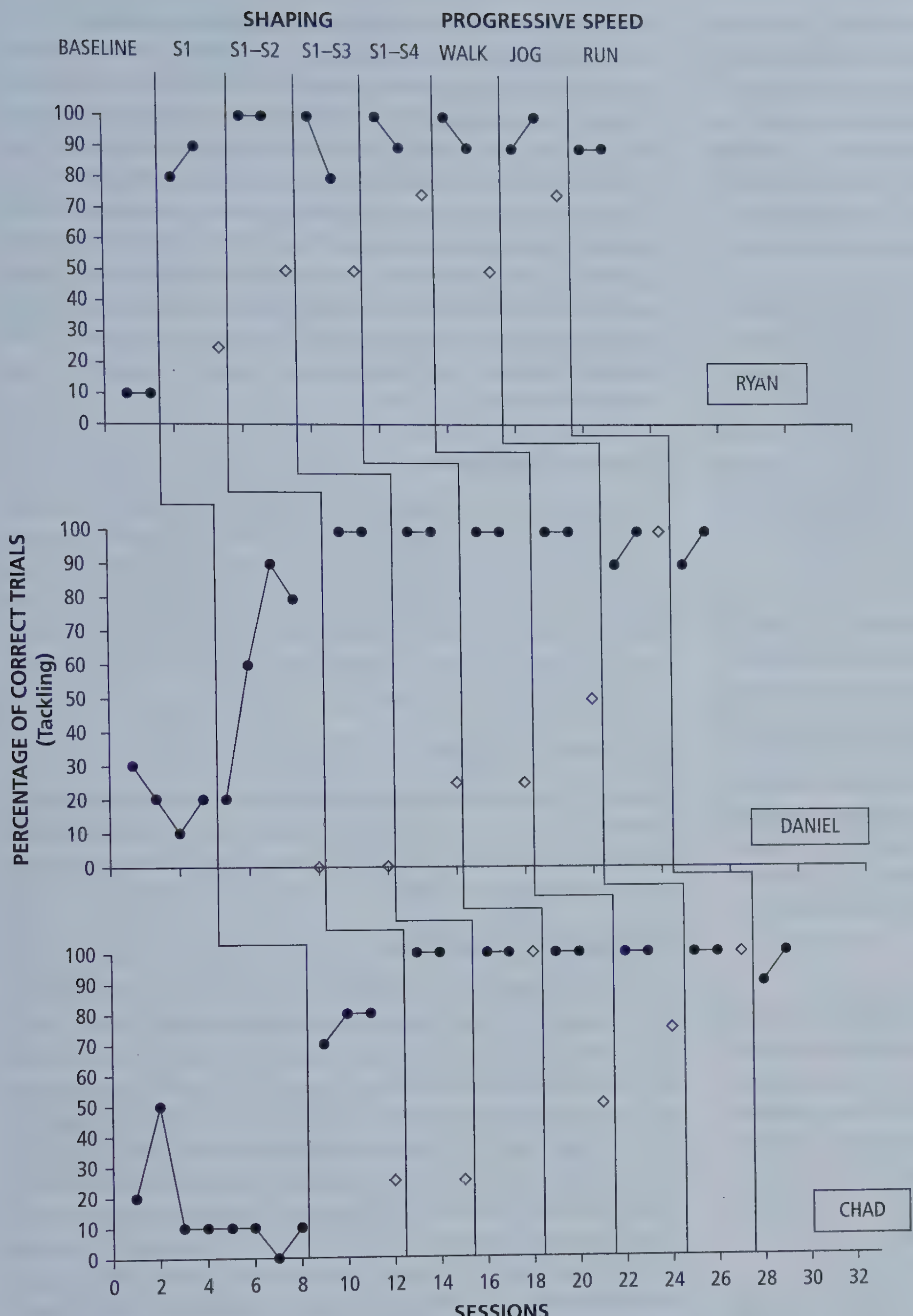


Figure 22.4 Percentage of correct tackling trials across all training phases. Open data points represent generalization probes of a complete tackle against a live ball carrier. (In the shaping procedure, S = skill.)

From "The Effects of Verbal Instruction and Shaping to Improve Tackling by High School Football Players," by A. H. Harrison and D. A. Pyles, 2013, *Journal of Applied Behavior Analysis, 46*(2), p. 521. Copyright by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

2. Progress toward the terminal behavior is not always linear. That is, the learner does not always proceed from one approximation to the next in a continuous, progressive, or errorless sequence. Progress may be erratic. If the behavior is too irregular (i.e., not resembling a closer approximation to the terminal behavior), an approximation may need to be further reduced, allowing for more reinforcement and progress. The skill of the practitioner in noting and reinforcing the next smallest approximation to the terminal behavior is critical to the success of shaping. If the practitioner fails to reinforce responses at the next approximation—because of inexperience, preoccupation with other tasks, or neglect—the occurrence of similar responses may be few and far between. If reinforcement for performance at a given approximation continues longer than necessary, progress toward the terminal behavior will be impeded.
3. Shaping requires *continuous* monitoring to detect subtle changes in the learner's performance, indicating the next approximation to the terminal behavior. Many practitioners—for example, teachers in busy or demanding classrooms—are not able to monitor behavior closely enough to note small changes. Consequently, shaping may be conducted inappropriately or at least inefficiently.
4. Shaping can be misapplied. Consider the child trying to obtain her father's attention by emitting low-volume demands (e.g., "Dad, I want some ice cream."). The father

does not attend to the initial calls. As the child's attempts become increasingly unsuccessful (i.e., not reinforced), she may become more determined to gain her father's attention. In doing so, the number and magnitude of the calls may increase due to extinction-induced variability (e.g., "DAD, I WANT ICE CREAM!"). After listening to the crescendo of vocal demands, the father eventually provides the ice cream. The next time, however, it may come to pass that the child states what she wants in an even louder voice before gaining her father's attention. In this scenario, the father has differentially reinforced an ever-rising level of attention-getting behavior and shaped higher levels of call-outs for ice cream. Using this example as a backdrop, Skinner (1953) anticipated the result of such a scenario: "Differential reinforcement applied by a preoccupied or negligent parent is very close to the procedure we should adopt if we were given the task of conditioning a child to be annoying" (p. 98).

5. Harmful behavior can be shaped. For example, Rasey and Iversen (1993) showed that differential reinforcement shaped a rat's behavior to the point that the rat fell from the edge of a platform. By differentially reinforcing the rat for sticking its nose farther over the edge for food, the rat eventually fell off the ledge.[4] It is not difficult to speculate how adolescent games such as *Dare and Double Dare* and videos that capitalize on persons receiving differential reinforcement from peers or "likes" and encouraging comments on social media for increasingly higher levels of risk taking can lead to dangerous—and sometimes tragic—behavioral outcomes.

Shaping Versus Stimulus Fading

Shaping and stimulus fading change behavior gradually, albeit in much different ways. In shaping, the antecedent stimulus for the terminal behavior stays the same, while the response progressively becomes more differentiated. In stimulus fading, the opposite occurs: The antecedent stimulus changes gradually, while the response stays essentially the same.

INCREASING SHAPING EFFICIENCY

In addition to showing how behaviors are shaped across and within response topographies, the Isaacs and colleagues (1960) study illustrates another aspect of shaping—efficiency. During the early stages of the program, the psychologist waited for the next approximation of the behavior to appear before delivering the reinforcer. Since waiting can be time-consuming and wasteful, Isaacs and colleagues improved efficiency by using a verbal response prompt, "Say gum," after the sixth training session. Presumably, if the psychologist had not used a verbal prompt, several additional sessions would have been necessary before a successful outcome was achieved. Shaping may become more efficient by using a discriminative stimulus, physical guidance, imitative prompts, and/or a percentile schedule.

Discriminative Stimulus

A discriminative stimulus (S^D) may be combined with shaping. For example, when attempting to shape hand shaking as a greeting skill for an adult with developmental disabilities, the teacher might say, "Frank, hold out your arm." Scott, Scott, and Goldwater (1997) used a discriminative stimulus in the form of a vocal prompt ("Reach!") as a university-level pole-vaulter ran a path to plant his pole in the vault box. The prompt was designed to focus the vaulter's attention on extending his arms prior to take-off in the vault.

Kazdin (2013) suggested that priming a response using any of a variety of prompting mechanisms (S^Ds) can be helpful, especially if the individual's repertoire is weak and the likelihood of distinguishable successive approximations is low. "Even if the responses are not in the client's repertoire of responses, the priming procedure can initiate early response components and facilitate shaping" (p. 277).

Physical Guidance

The second way to enhance shaping is through physical guidance. In the hand shaking example with Frank, cited earlier, the teacher might tap his arm manually, assist Frank with holding out his arm, move his arm forward with light pressure in a hand shaking motion, or hold his arm firmly through the full motion. Any physical guidance prompt used to enhance or speed up shaping should later be faded.

Imitative Prompts

Using an imitative prompt to demonstrate arm extension (e.g., "Frank, hold out your arm like this.") may improve shaping's efficiency. As with physical guidance, any imitative prompt should ultimately be faded.

Percentile Schedules

Because shaping requires a skilled observer to (a) detect behavior changes, (b) apply differential reinforcement to ever-closer approximations to a terminal behavior, and (c) place non-approximations on extinction, shaping has been characterized as more "art than science" (Lattel & Neef, 1996). Galbicka (1994) stated that practitioners would benefit from moving from a *qualitative*-based, observer-detecting standard for determining if the next response in a shaping program should be reinforced to a *quantitative* standard. In his view, a percentile schedule enables more precise shaping procedures, allows implementation by multiple trainers across settings, reduces the time it takes to complete a shaping sequence, and quantitatively moves closer to Skinner's conception of shaping.

> Percentile schedules disassemble the process of shaping into its constituent components, translate those components into simple, mathematical statements, and then use these equations, with parameters specified by the experimenter or trainer, to determine what presently constitutes a criterional response and should therefore be reinforced. (p. 740)

Using a mathematical equation ($i.e., k=(m+1)(1-w)$) and a computer software program, a percentile schedule continually calculates and adjusts the criterion for reinforcement based on recent observations.

> In this equation, w denotes the density of reinforcement, and m is a fixed number of recent observations. The k

> parameter specifies what response value out of *m* most recently observed response values an upcoming response must exceed to satisfy the criterion for reinforcement. For example, with a *w* value of .5, a response meeting the criterion for reinforcement for a given observation will be observed half the time, offering a mixture of reinforcement and extinction of responses. An *m* value of 10 means that the 10 most recent observations are ranked according to their ordinal value from least to most, keeping the shaping procedure in touch with the individual's current repertoire. Solving the equation using these values gives a *k* value of 5.5. In most cases, it is easier to implement the percentile schedule by rounding the *k* value to a whole number. In the example, rounding *k* to 5 keeps the reinforcement density at approximately half of all responses and makes the criterion for reinforcement slightly less stringent than a larger *k* value. The *k* value of 5 denotes the observation ranked fifth among the 10 most recently observed and ranked observations is the value the current observation must exceed to meet the criterion for reinforcement. (Athens, et al., 2007, pp. 475–476)

Several applied studies have evaluated percentile schedules with respect to shaping smoking cessation (Lamb, Kirby, Morral, Galbicka & Iguchi, 2010; Lamb, Morral, Kirby, Iguchi, & Galbicka, 2004), increased duration of eye contact (Hall, Maynes & Reiss, 2009), and increased engagement with academic tasks (Athens et al., 2007). In a parametric study of the effectiveness of percentile schedules in shaping academic task engagement, Athens and colleagues (2007) compared three values of *m* (5, 10, and 20). Their findings indicated that the percentile schedule was most effective when a relatively large number of previous observations ($N = 20$) were taken into account. In discussing the limitations of their study and comparing findings with those of other studies, Athens et al. noted that while "the procedure appears to be promising," the "potential utility of a percentile schedule for clinicians and applied researchers remains to be seen" (p. 486).

We agree with Athens et al. to the extent that a percentile schedule is encouraging. However, for greater utility in applied settings, practitioners would need to be well versed in using their recommended mathematical equation and be proficient with using the computer software program in "real time." Also, whether a percentile schedule increases shaping's efficiency over conventional applications remains an empirical question. Additional experimental studies that examine efficiency as a dependent variable will contribute to establishing percentile schedules as a viable alternative to conventional applications.

CLICKER TRAINING

Pryor (1999) described **clicker training** as a science-based system for progressively shaping new behavior using positive reinforcement and extinction. The clicker is a handheld device that produces a clicking sound when a metal tab is pressed. Reinforcement is paired with the sound of the clicker so that the sound becomes a conditioned reinforcer. Initially used to shape dolphin behavior without physical guidance (Pryor & Norris, 1991), clicker training was later used with other animals (e.g., cats and horses) and ultimately with humans to shape such complex behaviors as airplane pilot skills (M. Pryor, 2005). Other researchers and practitioners (e.g., Poling et al., 2011a) have used clicker training as part of their overall shaping program to teach African giant pouched rats how to detect land mines.

> Clicker trainers focus on building behavior, not stopping behavior. Instead of yelling at the dog for jumping up, you click it for sitting. Instead of kicking the horse to make it go, you click it for walking. Then, click by click, you "shape" longer sits, or more walking, until you have the final results you want. Once the behavior is learned, you keep it going with praise and approval and save the clicker and treats for the next new thing you want to train. (Pryor, 2002, p. 1)

Figure 22.5 provides 15 tips for getting started with clicker training (Pryor, 2002).

EMERGING APPLICATIONS OF SHAPING

At least three emerging applications of shaping can be considered: using computers to teach shaping skills, combining shaping procedures with robotics engineering, and applying telemedicine technology to implement shaping and differential reinforcement programs in hard-to-access urban, rural, or remote settings.

Using Computers to Teach Shaping

Several options are currently available to teach shaping skills using computers, or a combination of specialized software with computer-based applications. For example, Sniffy the Virtual Rat is a computer-based, digitized animation of a white rat within a Skinner box. Students practice basic shaping on their own personal computers using this virtual simulation.

Acknowledging the practical limitations of teaching behavior analysis principles (e.g., shaping) interactively to a large student body, Shimoff and Catania (1995) developed and refined computer simulations to teach the concepts of shaping.[5] In their simulation *The Shaping Game*, four levels of difficulty (easy, medium, hard, and very hard) were programmed over a series of refinements. Beginning with easier tasks and advancing to the very hard level, students are taught to detect successive approximations to bar-pressing behaviors that should be reinforced, and those that should not. In still later versions, students are expected to shape the movement of the animated rat from one side of a Skinner box to another. Shimoff and Catania made a case for computer simulations when they stated, "Computers can be important and effective tools for establishing sophisticated behavior that appears to be contingency shaped even in classes that do not provide real (unsimulated) laboratory experience" (pp. 315–316).

Martin and Pear (2003) contended that computers might be used to shape some dimensions of behavior (e.g., topography) more quickly than humans because a computer can be calibrated

Figure 22.5 Pryor's 15 tips for getting started with the clicker.

1. Push and release the springy end of the clicker, making a two-toned click. Then treat. Keep the treats small. Use a delicious treat at first: for a dog or cat, little cubes of roast chicken, not a lump of kibble.
2. Click DURING the desired behavior, not after it is completed. The timing of the click is crucial. Don't be dismayed if your pet stops the behavior when it hears the click. The click ends the behavior. Give the treat after that; the timing of the treat is not important.
3. Click when your dog or other pet does something you like. Begin with something easy that the pet is likely to do on its own (ideas: sit, come toward you, touch your hand with its nose, lift a foot, touch and follow a target object such as a pencil or a spoon).
4. Click once (in-out). If you want to express special enthusiasm, increase the number of treats, not the number of clicks.
5. Keep practice sessions short. Much more is learned in three sessions of five minutes each than in an hour of boring repetition. You can get dramatic results, and teach your pet many new things, by fitting a few clicks a day here and there in your normal routine.
6. Fix bad behavior by clicking good behavior. Click the puppy for relieving itself in the proper spot. Click for paws on the ground, not on the visitors. Instead of scolding for making noise, click for silence. Cure leash-pulling by clicking and treating those moments when the leash happens to go slack.
7. Click for voluntary (or accidental) movements toward your goal. You may coax or lure the animal into a movement or position, but don't push, pull, or hold it. Let the animal discover how to do the behavior on its own. If you need a leash for safety's sake, loop it over your shoulder or tie it to your belt.
8. Don't wait for the "whole picture" or the perfect behavior. Click and treat for small movements in the right direction. You want the dog to sit, and it starts to crouch in back: Click. You want it to come when called, and it takes a few steps your way: Click.
9. Keep raising your goal. As soon as you have a good response—when a dog, for example, is voluntarily lying down, coming toward you, or sitting repeatedly—start asking for more. Wait a few beats, until the dog stays down a little longer, comes a little further, sits a little faster. Then click. This is called "shaping" a behavior.
10. When your animal has learned to do something for clicks, it will begin showing you the behavior spontaneously, trying to get you to click. Now is the time to begin offering a cue, such as a word or a hand signal. Start clicking for that behavior if it happens during or after the cue. Start ignoring that behavior when the cue wasn't given.
11. Don't order the animal around; clicker training is not command-based. If your pet does not respond to a cue, it is not disobeying; it just hasn't learned the cue completely. Find more ways to cue it and click it for the desired behavior. Try working in a quieter, less distracting place for a while. If you have more than one pet, separate them for training, and let them take turns.
12. Carry a clicker and "catch" cute behaviors like cocking the head, chasing the tail, or holding up one foot. You can click for many different behaviors, whenever you happen to notice them, without confusing your pet.
13. If you get mad, put the clicker away. Don't mix scoldings, leash-jerking, and correction training with clicker training; you will lose the animal's confidence in the clicker and perhaps in you.
14. If you are not making progress with a particular behavior, you are probably clicking too late. Accurate timing is important. Get someone else to watch you, and perhaps to click for you, a few times.
15. Above all, have fun. Clicker-training is a wonderful way to enrich your relationship with any learner.

From *Click to win! Clicker training for the show ring* (Appendix A) by K. Pryor, 2002. Copyright © 2002 by Karen Pryor. Waltham, MA: Sunshine Books. Reprinted by permission.

From K. Pryor & L. Chamberlain, 2016, https://www.clickertraining.com/files/the-modern-principles-of-shaping-SF-edits.pdf. Copyright © 2016 by Karen Pryor. Reprinted with permission.

and programmed with specific decision rules for delivering reinforcement. They further suggested that in medical application cases (e.g., amputees, stroke victims), minute muscle movements may go undetected by a human observer, but a microprocessor could be calibrated to detect shapeable responses. In their view, "Computers . . . are both accurate and fast and may therefore be useful in answering fundamental questions concerning which shaping procedures are most effective . . . computers may be able to shape at least some kinds of behavior as effectively as humans" (p. 133).

Combining Shaping with Robotics Engineering

Dorigo and Colombetti (1998) and Saksida, Raymond, and Touretzky (1997) have explored how shaping can be applied to robot training. Essentially, shaping is used as one method of programming robots to progress through initial, intermediate, and final goals to achieve more complicated sets of commands. According to Savage (1998):

> Given that shaping is a significant determinant of how a wide variety of organisms adapt to the changing

> circumstances of the real world, then shaping may, indeed, have potential as a model for real-world robotic interactions. However, the success of this strategy depends on a clear understanding of the principles of biological shaping on the part of the roboticists, and the effective robotic implementation of these principles. (p. 321)

Although Dorigo and Colombetti (1998) and Saksida and colleagues (1997) concede the difficulty of applying successive approximations to robots, this line of inquiry, combined with emerging knowledge on artificial intelligence, offers exciting prospects for the future.

Applying Telemedicine Technology to Implement Shaping and Differential Reinforcement Programs in Hard-to-Access Urban, Rural, or Remote Settings

Telemedicine technologies involve connecting host-institution training centers with off-site rural or remote settings through high-speed, Internet-based, interactive audio-video links. Each site—host and remote—is equipped with the appropriate computers, cameras, monitors, and routers to enable instantaneous two-way communications between trainers at the host site and trainees at the remote site.

These technologies, originally developed to provide consultation and assessment services to families living in rural or remote locations (Barretto, Wacker, Harding, Lee, & Berg, 2006), have been expanded to provide additional services in the form of functional analyses and interventions, including differential reinforcement procedures (Peterson, Volkert, & Zeleny, 2015). Not only have telemedicine/telehealth technologies been shown to produce effective results, but also the added cost savings with respect to travel time and efficiency of implementation bode well for wider implementation in emerging rural/remote locations in the future.

Over the past decade, new and exciting applications have been explored to enhance the human condition and to increase safety and improve public welfare. Applications that combine shaping with discrimination training have been reported in several general areas: detecting land mines (Edwards, Cox, Weetjens, Tewelde, & Poling, 2015; Poling et al., 2011a); uncovering contraband (Jezierski et al., 2014); searching for, and locating, humans trapped in simulated collapsed buildings (Edwards, La Londe, Cox, Weetjens, & Poling, 2016; La Londe et al., 2015); and detecting tuberculosis or necrotic tissue (Poling et al., 2011b; Poling et al., 2017). Poling et al. (2017), for example, reported that trained African giant pouched rats detected 2000 additional active tuberculous cases compared to cases that reported tuberculous using only traditional microscopic analysis for diagnosis.

Likewise, in the case of land mine detection, Poling et al. (2011a) incorporated shaping, clicker training, and differential reinforcement of African giant pouched rats, who, after training, successfully searched over 93,000 m^2 of land (i.e., approximately 23 acres),[6] locating upwards of 100 mines and other explosives. The rats were 100% accurate in discriminating buried land mines from other materials buried beneath the soil surface, a finding replicated by Edwards et al. (2015) with remarkable consistency.

It is worth noting that these field-based efforts provide stunning examples of how shaping can be combined with other procedures (namely, clicker and discrimination training) to help communities plagued by disease or the ravages of war to reclaim healthier and more productive lives.

SHAPING GUIDELINES

The practitioner should consider various factors before deciding to use shaping. First, assess the terminal behavior and the available resources. For example, assume that a fifth-grade teacher is interested in increasing the number of math problems performed by a student with learning disabilities. The student completes 5 problems per math period, with a range of 0 to 10 problems. If the student has the ability to check her performance independently at the end of the class period, shaping could be implemented whereby the number of correctly completed math problems per period would be differentially reinforced. Reinforcement would be presented for 5, 7, 9, and progressively more problems answered correctly per session.

Second, since shaping may require multiple approximations and additional trained personnel, the analyst should assess whether time constraints, staff issues, or lack of resources militate against the use of shaping.

Third, some behaviors seem to preclude the use of shaping. For instance, if the manager of a community-based organization (e.g., *Toastmasters International*) is interested in increasing the public-speaking repertoires of participants, prompting, modeling, or peer tutoring might be more efficient than shaping. That is, telling or showing the participant how to use a gesture, voice inflection, or a metaphor would likely be much faster than attempting to shape each of these presentation skills alone.

Finally, Pryor's original *Ten Laws of Shaping* (1999) have evolved over time into what is now termed *The Modern Principles of Shaping*. These principles can assist the practitioner in implementing shaping in applied settings (see Figure 22.6).

Select the Terminal Behavior

Practitioners often work with learners having multiple behavior challenges. Consequently, they must identify the highest priority behavior to change as quickly as possible. The ultimate criterion in this decision is the individual's expected independence after an intervention—that is, the likelihood of her attaining additional reinforcers from the environment (Brown, McDonnell, Snell, & Brown, 2016). For example, if a student frequently roams the classroom poking other students, taking their papers, and verbally harassing them, a shaping procedure might best begin with a behavior that is incompatible with roaming the room because of the utility it would have for the individual and the other students. In addition, if in-seat behavior is developed, the staff with whom the student interacts will likely notice and reinforce it. In this case, shaping should differentially reinforce longer durations of in-seat behavior.

Figure 22.6 Pryor's *The Modern Principles of Shaping.*

1. **Be prepared before you start.** Be ready to click/treat *immediately* when the training session begins. When shaping a new behavior, be ready to capture the very first tiny inclination the animal gives you toward your goal behavior. This is especially true when working with a prop such as a target stick or a mat on the ground.
2. **Ensure success at each step.** Break behavior down into small enough pieces that the learner always has a realistic chance to earn a reinforcer.
3. **Train one criterion at a time.** Shaping for two criteria or aspects of a behavior simultaneously can be very confusing. One click should not mean two different criteria.
4. **Relax criteria when something changes.** When introducing a new criterion or aspect of the skill, temporarily relax the old criteria for previously mastered skills.
5. **If one door closes, find another.** If a particular shaping procedure is not progressing, try another way.
6. **Keep training sessions continuous.** The animal should be continuously engaged in the learning process throughout the session. He should be working the entire time, except for the moment he's consuming/enjoying his reinforcer. This also means keeping a high rate of reinforcement.
7. **Go back to kindergarten, if necessary.** If a behavior deteriorates, quickly revisit the last successful approximation or two so that the animal can easily earn reinforcers.
8. **Keep your attention on your learner.** Interrupting a training session gratuitously by taking a phone call, chatting, or doing something else that can wait often causes learners to lose momentum and get frustrated by the lack of information. If you need to take a break, give the animal a "goodbye present," such as a small handful of treats.
9. **Stay ahead of your learner.** Be prepared to "skip ahead" in your shaping plan if your learner makes a sudden leap.
10. **Quit while you're ahead.** End each session with something the learner finds reinforcing. If possible, end a session on a strong behavioral response, but, at any rate, try to end with your learner still eager to go on.

From K. Pryor & L. Chamberlain, 2016, https://www.clickertraining.com/files/the-modern-principles-of-shaping-SF-edits.pdf.

It is also important to define the terminal behavior precisely. For example, a behavior analyst might want to shape the appropriate sitting behavior of an individual with severe intellectual disabilities. To do so, sitting might be defined as being located upright in the chair, facing the front of the room with buttocks against the bottom of the chair and back against the chair rest during a 15-min morning activity. Using this definition, the analyst can determine when the behavior is achieved, as well as what does not constitute the behavior—an important discrimination if shaping is to be conducted efficiently (e.g., the student might be half in and half out of the seat, or the student might be in the seat but turned toward the back of the room).

Determine the Criterion for Success

After identifying the terminal behavior, the practitioner should specify the criterion for success. Here, the practitioner decides how accurate, fast, intense, or durable the behavior must be before it can be considered shaped. Several measures can be applied to establish a criterion of success. Some of the more common include rate, magnitude, and duration. Depending on the terminal behavior, any or all of these dimensions can be used to assess achievement. For instance, a special education teacher might shape rate performance in solving mathematics calculations problems that would proceed from the student's current level of performance (say, half a problem per minute) to a higher rate (e.g., four problems per minute).

Norms for success can be determined by measuring the behavior in a similar peer group, consulting established norms in the literature, or complying with international standards for accuracy. For instance, norms and guidelines for reading rate by grade level (Kame'enui, 2002), physical education skills for children by age group (President's Council on Physical Fitness, 2016), homework guidelines per grade level (Heron, Hippler, & Tincani, 2003), or accepted standards for land mine detection (Poling, 2011a; Edwards et al., 2015) can be referenced to validate progress toward, or the achievement of, the terminal behavior.

In the earlier illustration, the criterion for success could be the student's sitting in her seat appropriately 90% of the time during the morning activity for 5 consecutive days. In this example, two criteria are specified: the percentage of acceptable sitting behavior per session (i.e., 90%) and the number of days that the criterion must be met to establish the behavior (i.e., 5 consecutive days). Poling and colleagues (2011a) used multiple criteria for determining success with respect to the accurate performance trials per day, and consecutive days at increasingly higher performance levels.

Assess the Response Repertoire

There is a two-fold reason for assessing response repertoires. First, having established the desired terminal behavior, it is important to ascertain the "behavioral gap" between the current level of behavior and the desired target behavior. In our respiratory therapy example cited earlier, that gap constituted 1500 ml of volume difference (i.e., goal was 1750 ml; present level was 250 ml). Second, assessing the response repertoire may help to project the successive approximations that might be emitted in the shaping sequence. When approximation behaviors

are known—or anticipated—the practitioner is in a better position to observe and reinforce an approximation when it occurs. However, the practitioner must recognize that projected approximations are guesstimates of behaviors. In actuality, the learner may emit behaviors that are merely approximations of these projections. The practitioner must be able to make the professional judgment that the previously emitted behavior either is, or is not, a closer approximation to the terminal behavior than the behaviors that occurred and were reinforced in the past. Galbicka (1994) provides a guideline for the behavior analyst with respect to assessing an existing response repertoire.

> The successful shaper must carefully ascertain characteristics of an individual's present response repertoire, explicitly define characteristics the final behavior will have at the end of training, and plot a course between reinforcement and extinction that will bring the right responses along at the right time, fostering the final behavioral sequence while never losing responding altogether (p. 739).

The relevant approximations across or within a response repertoire can be analyzed in several ways. First, practitioners can consult experts in the field to determine their views on the proper sequence of approximations for a given behavior (Brown et al., 2016). For example, teachers who have taught three-digit multiplication for several years can be consulted to learn their experiences with prerequisite behaviors for performing this type of calculation task. Second, as stated earlier, normative data from published studies may be used to provide an estimate of the approximations involved. Third, a digital recording can be used to analyze the component behaviors. A slow-motion digital replay would help the practitioner to see motion that might be undetectable in an initial faster-moving live performance, but become detectable as the practitioner becomes skilled with observing it. Finally, the practitioner can perform the target behavior himself, carefully noting the discrete behavioral components as he produces them.

The ultimate determination of the approximations, the approximation order, the duration of time that reinforcement should be delivered at a given approximation, and the criteria for skipping or repeating approximations rests with the skill and the judgment of the practitioner. Ultimately, the learner's performance should dictate when approximation size should be increased, maintained, or decreased, making consistent and vigilant monitoring by the practitioner essential.

Identifying Behaviors to Reinforce

At the outset of any shaping session, the shaper must observe the participant closely to detect any movement, slight though it might be, in the direction of the desired terminal behavior. At the instant that this behavior occurs, reinforcement is delivered. So, if an occupational therapist was shaping the terminal behavior of a toddler to eat a bowl of potatoes with a spoon, any hand or finger movement, even a twitch, would be reinforced. Subsequently, successive hand or finger movements toward the spoon would be reinforced. Thereafter, movements that successively touched the spoon, grasped the spoon, raised the spoon, directed the lifted spoon toward her mouth, and so on, would be reinforced, whereas other behaviors would undergo extinction. Conversely, the analyst would *not* set an initial criterion of raising the spoon out of the bowl. To set the behavior criterion at this level would be counterproductive, as raising-the-spoon behavior may not be in the child's present repertoire.

Eliminate Interfering or Extraneous Stimuli

Eliminating distractions during shaping enhances the process. For example, if a parent is interested in shaping one dimension of her young daughter's dressing behavior (e.g., the rate of dressing before school), but decides to begin the shaping procedure in her bedroom, where cartoons are playing on a morning television show, the shaping program may not be successful. The cartoons will compete for the daughter's attention. It would be more efficient to choose a time and a location where such sources of distraction can be reduced or eliminated.

Proceed Gradually

The importance of proceeding toward the terminal goal gradually cannot be overemphasized. The practitioner should anticipate changes in the rate of progress and be prepared to go from approximation to approximation as the learner's behavior dictates. Each new occurrence of a successive approximation to the terminal behavior must be detected and reinforced. If not, shaping will be, at worst, unsuccessful or, at best, haphazard, requiring much more time. Furthermore, the occurrence of a behavior at a given approximation does not mean the next response that approximates the terminal behavior will be immediately or accurately produced.

The practitioner must also be aware that many trials may be required at a given approximation before the subject can advance to the next approximation. Figure 22.7 illustrates this point well. Scott and colleagues (1997) found that the percentage of correct arm extensions by the pole-vaulter decreased initially as the height of the bar was raised. Whereas a 90% correct arm extension was noted at the initial height, the vaulter's correct performance typically dropped to approximately 70% each time the bar was raised. After successive attempts at the new height, the 90% correct criterion was re-established.

Limit the Number of Approximations at Each Level

Just as it is important to proceed gradually from approximation to approximation, it is equally important to ensure that progress is not impeded by offering too many trials at a given approximation. This may cause the behavior to become firmly established, and because that approximation has to be extinguished before progress can begin again (Catania, 2013), the more frequently reinforcement is delivered at a given approximation, the longer the process can take. In general, if the learner is progressing steadily, reinforcement is being delivered at the correct pace. If numerous mistakes are being made, or behavior ceases altogether, the criterion for reinforcement has been raised too quickly. Finally, if the performance stabilizes at a certain level, the pace of shaping is too slow. Practitioners may experience all three of these conditions while shaping. Therefore, they must be vigilant and adjust the next successive approximation and reinforcement schedule as necessary.

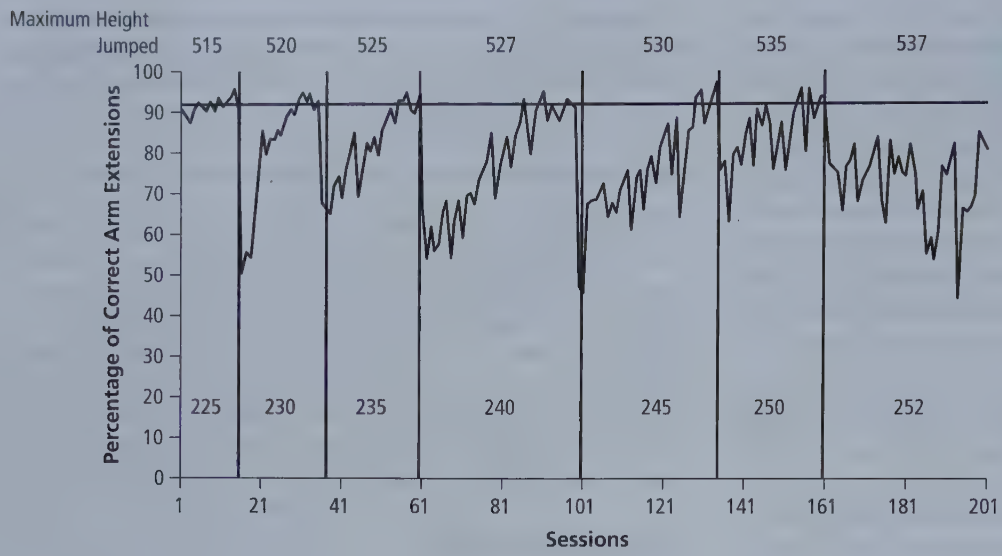


Figure 22.7 Percentage of correct arm extensions at seven different heights. The numbers contained within each criterion condition refer to the height of the photoelectric beam.

From "A Performance Improvement Program for an International-Level Track and Field Athlete," by D. Scott, L. M. Scott, and B. Goldwater, 1997, *Journal of Applied Behavior Analysis, 30*, p. 575. Copyright by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

Continue to Reinforce When the Terminal Behavior Is Achieved

When the terminal behavior is demonstrated, and reinforced, it is necessary to continue to reinforce it over time; otherwise, the behavior will extinguish, and performance will return to a lower level. Reinforcement must continue until the criterion for success is achieved, and a maintenance schedule of reinforcement is established.

LEARNING TO SHAPE

Reading the narrative, illustrations, rules, and guidelines, such as those presented in this chapter, can help an emerging analyst become a skilled behavior shaper. Such reading is necessary, but not sufficient, with respect to gaining proficiency as a "shaper." The best way to learn how to shape behavior proficiently is to practice shaping behavior! So, how does one go about practicing shaping behavior? Let's explore a few options that can be arranged whether you prefer to practice by yourself, with a partner, or in a group. Whatever the arrangement, however, seek the advice and direction of a skilled behavior shaper who can provide directive feedback and reinforcement during your training.

By Yourself

You could practice trying to shape a simple terminal behavior with an animal (e.g., teaching a rat to press a bar, a pigeon to peck a disc, or a dog to follow a "sit" command). The simpler the terminal behavior and the more "controlled" the setting, the better (e.g., a lab, a quiet room away from distractions). Previewing a well-conducted shaping session on a YouTube video that resembles the behavior you are trying to shape, or consulting with a skilled colleague using Skype or FaceTime, might expedite the learning process for you.

Alternatively, where access to animals, controlled laboratories, or cooperative teachers is limited, consider using a simulation program such as Sniffy the Virtual Rat (Alloway, Wilson, & Graham, 2005). Sniffy, as mentioned earlier, is a computer-based, digitized, virtual animation of a white rat that can be accessed on a personal computer, which allows one to practice basic shaping at any convenient time or location.

With a Partner

Choose a friend or colleague and play "*PORTL*," an acronym for *Portable Operant Research and Teaching Laboratory* (Hunter & Rosales-Ruiz, 2018). Box 22.1 outlines the rationale for this exercise, and explains how to play PORTL with a partner for maximum benefit.[7]

With a Group

Keenan and Dillenburger (2000) described a classroom activity that teaches students shaping principles. Keenan and Dillenburger's "The Shaping Game," though sharing an identical name with Shimoff and Catania's (1995) computer simulation, is an *in vivo* experience shared by a group and it has a number of variations. The principal manner in which this game is played is to have a volunteer student initially exit the classroom temporarily while the remaining class members, guided by the instructor, reach consensus on the target behavior to shape when the volunteer re-enters the classroom. Target behaviors can be simple—walking to the front of the room and turning off a light switch—or be more elaborate movements (e.g., walking to the front of the room, standing on one foot while touching an elbow with the opposite hand). A volunteer "shaper" is also chosen.[8] Upon re-entry, the shaper clicks and then reinforces the first behavior that approximates the terminal behavior. Other behaviors or movements are placed on extinction. The Shaping Game concludes when the volunteer performs the terminal behavior targeted by the class. According to Keenan and Dillenburger the upshot of The Shaping Game is to:

1. Teach the basic principles involved in shaping.
2. Teach about the power of contingencies of reinforcement.
3. Demonstrate the operation of stimulus control.
4. Demonstrate the value and difficulty of continuously monitoring the behavior of a single organism.
5. Demonstrate how questionnaires produce an incomplete account of behavior (p. 22).

BOX 22.1

Learn Shaping by Playing PORTL

PORTL (Portable Operant Research and Teaching Lab) is a second-generation table-top shaping game, originally developed by canine trainer Kay Laurence and now extrapolated to human participants by Jose Rosales-Ruiz and his graduate students at the University of North Texas.

Rationale for Playing PORTL

PORTL teaches participants to shape behavior entirely through delivering reinforcement and changing environmental arrangements. Early training exercises help individuals practice shaping basics—observing behavior, clicking a clicker at the next successive approximation movement of their partner, reinforcing simple behaviors consistently, and collecting and recording data collection—without the benefit of verbal instructions, models, gestures, or prompts. During intermediate training, individuals continue to learn additional concepts, including cueing and stimulus control, concept formation, chaining, schedules of reinforcement, superstitious behavior, extinction, and resurgence, to name a few. Ultimately, individuals gain experiences being the teacher *and* the learner, which helps to build understanding and empathy as the learner relates to the challenges the game presents to his or her students.

How Is PORTL Played?

PORTL is played between two individuals, the teacher and the learner, using a collection of small objects, a clicker to "mark" behavior, and small tokens or blocks as reinforcers.

At the beginning of each exercise, the teacher presents the learner with a selection of objects. The learner is given only two instructions: "Your goal is to earn as many blocks as possible" and "Please interact with the objects."

When the learner performs the correct behavior or an approximation toward the correct behavior, the teacher clicks the clicker and delivers a contingent token. The learner places the token into a container and then resumes interacting with the objects. The teacher can rearrange the objects or add and remove objects to help guide the learner toward the goal behavior.

The teacher and learner take breaks periodically to fill out datasheets. The teacher records what just happened and what she plans to teach during the next training period. The learner records what he thinks he is learning and his perception of the exercise. Breaks can be taken after a criterion number of tokens have been delivered (e.g., 10 tokens) or after a certain amount of time (e.g., 60 seconds). Breaks allow the teacher to continually assess the learner's progress and to adjust the shaping plan.

After each exercise, the teacher and learner compare datasheets. In addition, the learner can provide valuable feedback about what the teacher did well, as well as insight regarding which parts of the exercise may have been confusing and why.

Adapted from M. Hunter and J. Rosales-Ruiz, 2018, *An introduction to PORTL: The Portable Operant Research and Teaching Lab*. https://www.artandscienceofanimaltraining.org/tools/portl-shaping-game/. Used by permission.

SUMMARY

Shaping Defined

1. Shaping consists of a three-part process whereby the analyst (a) detects a change in the learner's environment, (b) makes a discriminated judgment about whether that change is a progressively closer approximation to a terminal behavior of interest, and then (c) differentially reinforces that closer successive approximation. Shaping is defined as the differential reinforcement of successive approximations toward a terminal behavior.

2. A terminal behavior—the end product of shaping—can be claimed when the topography, rate, latency, duration, interresponse time, or magnitude of the target behavior reaches a predetermined criterion level.

3. Differential reinforcement refers to the act of presenting an unconditioned or conditioned reinforcer only to those emitted members of a response class that share a specified dimension or quality, while placing all other emitted response class members on extinction.

4. Differential reinforcement has three effects: Responses similar to those that have been reinforced are more likely to occur, responses resembling the unreinforced members are emitted less frequently (i.e., undergo extinction), and some of these emitted responses under extinction may be likely to resemble the next successive approximation to the terminal behavior and would therefore be targeted for reinforcement.

5. When differential reinforcement is applied consistently within a response class, its triple effect results in a new response class composed primarily of responses sharing the characteristics of the previously reinforced subclass. This emergence of a new response class is called response differentiation.

6. The gradually changing criterion for reinforcement during shaping results in a succession of new response classes, or successive approximations, each one closer in form to the terminal behavior than the response class it replaces.

7. Behavior can be shaped along measurable dimensions of topography (form), rate (number of responses per unit of time), latency (time between the onset of the antecedent stimulus and the occurrence of the behavior), duration (elapsed time from the onset of a response to its end point), interresponse time (time between responses), and magnitude (response strength or force).

Shaping Across and Within Response Topographies

8. Shaping behavior across different response topographies means that selected members of a response class are differentially reinforced, whereas members of other response classes are placed on extinction.
9. Shaping a behavior within a response topography means that the form of the behavior remains constant, but differential reinforcement is applied to another measurable dimension of the behavior (e.g., magnitude).
10. Shaping is a positive approach to teach new behavior and may be combined with other behavior change tactics.
11. Shaping has at least five limitations that practitioners should consider before applying it.
12. In shaping, the antecedent stimulus stays the same, while the response progressively becomes more differentiated. In stimulus fading, the response remains the same, while the antecedent stimulus changes gradually.

Increasing Shaping Efficiency

13. Shaping may become more efficient by using a discriminative stimulus, physical guidance, imitative prompts, and/or a percentile schedule. Any prompt that is introduced should be faded later.

Clicker Training

14. Clicker training is a science-based system for progressively shaping behavior using positive reinforcement and extinction.
15. The clicker, a handheld device that produces a clicking sound when a metal tab is pressed, provides the signal that when a behavior is being performed in the presence of the click, reinforcement follows.

Emerging Applications of Shaping

16. Using computers to teach shaping, combining shaping procedures with robotics engineering, and applying telemedicine technology for humanitarian purposes for populations in hard-to-access settings serve as three emerging applications of shaping.

Shaping Guidelines

17. Before deciding to use shaping, assess the nature of the behaviors to be learned, the staff requirements and resources available, and whether a different tactic altogether might be more efficient in terms of time and effort.
18. After the decision has been made to use a shaping procedure, the practitioner performs the following steps: select the terminal behavior, decide the criterion for success, analyze the response class, identify the first behavior to reinforce, eliminate interfering or extraneous stimuli, proceed gradually, limit the number of approximations at each level, continue reinforcement when the terminal behavior is achieved, and learn to shape behavior.

Learning to Shape

19. To learn to shape behavior, consider shaping a simple behavior, working with a colleague, practicing with a group, or using a computer-based simulator.

KEY TERMS

clicker training
differential reinforcement
extinction-induced variability
response differentiation
shaping
successive approximation

MULTIPLE-CHOICE QUESTIONS

1. Shaping involves:
   a. Reinforcing successive approximations to a terminal behavior
   b. Creating response differentiation between responses that are closer to the terminal behavior and those that are not
   c. Placing responses that are not a closer approximation of the terminal behavior on extinction
   d. All of these

   Hint: (See "Shaping defined")
2. Which of the following is an example of shaping an individual to swing a golf club correctly (i.e., take the appropriate backswing, swing the club down to the ball, and follow through with the swing)?
   a. Stand behind the person and physically help make the correct golf swing. Over time, slowly provide less physical assistance until the individual makes a correct swing independently.
   b. Start with verbal prompts to do each step. If this is not effective, provide a model. If this is not effective,

provide some physical assistance. If this is not effective, stand behind the person and help the person make the correct golf swing.

c. Provide positive reinforcement for any swing that resembles a golf swing. Then, provide reinforcement for closer and closer approximations of a correct golf swing and withhold reinforcement for poorer approximations of a correct golf swing.

d. None of these is an example of shaping.

Hint: (See "Shaping defined")

3. Response differentiation involves what two components?

a. Differential reinforcement and changing the criterion for reinforcement

b. Punishment and extinction

c. Prompting and differential reinforcement

d. Extinction and changing the criterion for reinforcement

Hint: (See "Shaping defined")

4. Ms. Anita Response wants to increase the number of words Sophie writes in her stories during creative writing time. She uses shaping to do so. What dimension of behavior is she shaping?

a. Amplitude

b. Frequency

c. Duration

d. Latency

Hint: (See "Shaping defined")

5. An example of shaping within a response topography is:

a. Increasing the number of math problems an individual needs to complete during math class

b. Increasing the duration of in-seat behavior during seatwork time

c. Increasing the force of a tennis swing

d. All of these are examples of shaping within a response topography

Hint: (See "Shaping across and within Response Topographies")

6. The text gives an example of shaping gum chewing, in which lip movement first received reinforcement, then lip movements with sound, then saying the word "gum," and finally, saying "Gum please." This is an example of:

a. Shaping within a response topography.

b. Shaping across different response topographies.

c. Shaping the magnitude of a response.

d. Shaping the duration of a response.

Hint: (See "Shaping across and within Response Topographies")

7. Shaping can be a slow, time-consuming process. What is one way to effectively speed up the process?

a. Add a punishment contingency for incorrect responses.

b. Each time you increase the criteria for reinforcement, do so by larger increments.

c. Model the desired response for the individual.

d. All of these are effective ways to make shaping more efficient.

Hint: (See "Increasing Shaping Efficiency")

8. When is adding a prompt to the shaping process something to consider?

a. When the individual is progressing through the shaping procedure quickly.

b. When the individual has mastered the terminal behavior.

c. When the individual's repertoire is weak and the likelihood of making a successive approximation is low.

d. When you have limited staff and resources to implement the shaping procedure.

Hint: (See "Increasing Shaping Efficiency")

9. What is meant by the statement that progress in shaping is rarely linear and often erratic?

a. Most individuals are unsuccessful in shaping new responses.

b. The amount of time it may take to shape a behavior is often unpredictable.

c. Sometimes individuals may exceed your criteria for reinforcement, making closer approximations than you had expected.

d. Sometimes individuals may get "stuck" on a response and have to "back up" to a previously-learned approximation before making process. This can slow the process down more than expected.

e. All of these

f. None of these

Hint: (See "Shaping Guidelines")

10. One of the guidelines for effective shaping is to carefully define the terminal behavior and the criteria for successfully engaging in the terminal behavior. Effective ways for deciding what are the criteria for success:

a. Read the research literature to see if there are published norms and guidelines available for the behavior you are trying to shape.

b. Observe competent peers to see how well they perform the task, and set your criteria according to their performance level.

c. Both of these are ways to determine the criteria for success.

d. Neither of these are appropriate ways to determine criteria for success.

Hint: (See "Shaping Guidelines")

ESSAY-TYPE QUESTIONS

1. Write the definition of shaping and provide an example of shaping that has not been discussed in class or presented in the textbook. When giving your example, analyze the response class and diagram the shaping process.

   Hint: (See "Shaping Defined")

2. What are the behavioral processes involved in the shaping process?

   Hint: (See "Shaping Defined")

3. What is the key difference between shaping across response topographies and shaping within a response topography?

   Hint: (See "Shaping across and within Different Response Topographies")

4. Shaping a new response can take a long time. What is one way you can increase the efficiency of shaping and speed up the process?

   Hint: (See "Increasing Shaping Efficiency")

5. What key steps are to be taken to implement the shaping process?

   Hint: (See "Shaping Guidelines")

NOTES

1. When applied to discrimination training, differential reinforcement requires (1) reinforcing responses in the presence of one stimulus condition (the S^D) and (2) placing responses in the absence of the S^D condition on extinction. Refer to Chapters 17 and 24 for this discussion.
2. Spirometers are not calibrated to show minute, incremental changes. They register units in larger whole numbers (500 ml, 1000 ml, etc.). The point remains that the therapist is actively (a) watching the spirometer to detect a change in breathing, (b) judging whether that change is a successive approximation higher than a previous attempt, and (c) reinforcing it if it is. Simultaneously, attempts that do not approximate the terminal behavior are placed on extinction.
3. Procedures for chaining a sequence of responses to a terminal outcome are described in Chapter 23.
4. The researcher provided a safety net so that the rat was not injured.
5. They also developed simulations for reinforcement schedules.
6. An American-size football field, end-zone-to-end-zone and sideline-to-sideline, is approximately 1.32 acres.
7. Accessing https://youtu.be/YSu7ghj4ib0 shows a video clip of PORTL being conducted.
8. Some variations have the entire class serve as the shaper by having them clap collectively when an approximation occurs.

CHAPTER 23

Chaining

LEARNING OBJECTIVES

- Define and give examples of behavior chains.
- Identify why behavior chains are important.
- Conduct a task analysis and select appropriate assessment methods for evaluating performance on task analysis.
- Identify and define different chaining procedures and determine when each would be selected for use.
- Describe how behavior chains can be interrupted.

At age 21, I was a newly employed counselor at a summer camp designed to provide educational, social, and recreational experiences to children with cerebral palsy. One of the campers—let's call him "Joe"—wore flapjack shoes that featured a mechanical upper tongue that clamped down to secure the shoes to his feet. I soon learned that Joe wore the flapjacks not as a fashion statement, but because he hadn't learned to tie his own shoelaces. A senior counselor challenged me to teach Joe to tie his shoes before the camp ended in approximately two weeks. I happily accepted the challenge, as an authentic Philadelphia cheesesteak sandwich hung in the balance.

I launched this project with my "go to" procedure—backward chaining. I used a practice jig that consisted of canvas cloth with holes at each border wrapped around a wooden board. I tied the laces completely, but loosely, and then held the jig out to Joe and said, "Tie your shoe." Joe snugged the laces tight, the last behavior in the chain. I reset the jig to my initial position, and repeated my direction to Joe, "Tie your shoe." Once Joe mastered the snugging task, the second behavior to teach was the next-to-the-last task in the shoe-tying sequence: drawing the looped laces down to the wooden template (before snugging them tight as before). And so the process continued until the first task in the sequence (i.e., crossing the laces) was completed. Each time Joe completed the chain, from performing the final step only in his very first trial to completing each step in sequence from the beginning, he contacted the natural reinforcer (a self-tied shoe) and social praise ("Good job, Joe!").

Practice occurred for about 10 to 15 minutes per day, three days per week. In the remaining days of the camp, Joe practiced the entire chain a few times with an Oxford-style shoe. However, on the last day of the camp, and sporting a new pair of sneakers, the senior counselor asked Joe to tie his shoe. Joe responded with the precision of a Swiss watchmaker. I reveled in Joe's success and enjoyed a delightful lunch to boot! Backward chaining chalked up another victory.

—Tim Heron

This chapter defines a behavior chain, provides a rationale for establishing behavior chains in applied settings, and discusses the importance of task analysis in behavior chain training. The chapter presents a procedures for constructing and validating task analyses, along with procedures for assessing individual mastery levels. Next, four behavior change methods—forward chaining, total-task chaining, backward chaining, and backward chaining with leap aheads—are addressed, followed by guidelines for deciding which behavior chain procedure to use in applied settings. Techniques for interrupting and breaking inappropriate chains are also provided. The chapter concludes with an examination of factors affecting the performance of a behavior chain.

BEHAVIOR CHAIN DEFINED

A **behavior chain** is a linked sequence of responses leading to a terminal outcome. Each response produces a stimulus change that functions as conditioned reinforcement for that response and as a discriminative stimulus (S^D) for the next response in the chain. Reinforcement for the last response in a chain maintains the reinforcing effectiveness of the stimulus changes produced by all previous responses in the chain.

Reynolds (1975) provided this laboratory example of a behavior chain (illustrated in Figure 23.1):

> Assume that a pigeon is presented with a blue light key (S_1). When the pigeon pecks the key (R_1), the blue light terminates, and a red light is activated. In the presence of the red light key (S_2), the pigeon presses a pedal (R_2) that turns the key to a yellow light. The yellow light (S_3) serves as an S^D to displace a bar (R_3), which terminates the yellow light, changing it to a green light (S_4). Finally, during the green light (S_4), pecks terminate the green light, and a grain-delivery mechanism is activated that dispenses an edible reinforcer. . . . Because each stimulus has a dual function, as discriminative stimulus and conditioner reinforcer, the links overlap. In fact, it is this dual function of the stimuli that holds the chain together. (pp. 59–60)

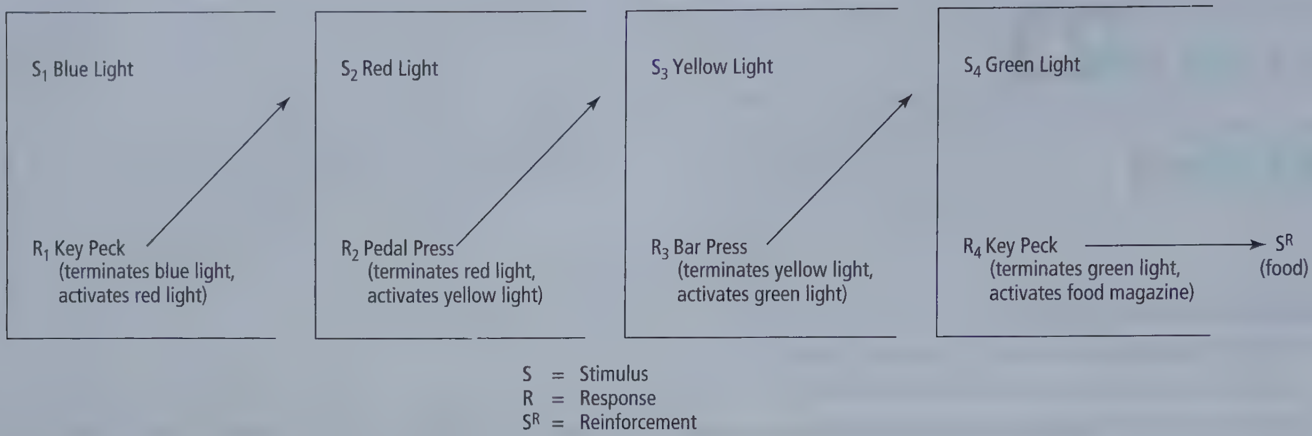


Figure 23.1 Illustration of a behavior chain consisting of four components.
Based on a chain described by Reynolds (1975, pp. 59–60).

As Figure 23.1 shows, this chain includes four responses (R_1, R_2, R_3, and R_4), with a specific stimulus condition (S_1, S_2, S_3, and S_4) associated with each response. The blue light (S_1) is an S^D to occasion the first response (R_1), a key peck that terminates the blue light and produces the onset of the red light (S_2). The red light serves as conditioned reinforcement for R_1 and as an S^D for R_2 (a pedal press). The pedal press (R_2) terminates the red light and produces the onset of the yellow light (S_3), and so on. The last response produces food, thus completing and maintaining the chain.

Figure 23.2 shows the links of a behavior chain using a classroom example. In preparing a class of preschool students for recess, the teacher might say to one student, "Please put on your coat." The teacher's statement would serve as the S^D (S_1) evoking the first response (R_1) in the chain, securing the coat in hand from the closet. That response in the presence of the teacher's statement terminates the teacher's statement and produces the onset of the coat in the student's hands (S_2). The coat in the student's hands serves as conditioned reinforcement for obtaining the coat from the closet (R_1) and as an S^D (S_2) for putting one arm through a sleeve (R_2). That response, putting one arm through the sleeve in the presence of the coat in both hands, terminates the stimulus condition of the coat in the student's hands and produces the onset of one arm in a sleeve and one arm out (S_3), a stimulus change that serves as conditioned reinforcement for putting one arm through a sleeve and as an S^D (S_3) for placing the second arm through the other sleeve (R_3). That response terminates the condition of one arm in a sleeve and one arm out, and produces the S^D (S_4) of the coat being completely on. The coat being fully on serves as a conditioned reinforcer for putting the second arm through the other sleeve and as an S^D for zipping it (R_4). Finally, zipping up the coat completes the chain and produces teacher praise, and the natural consequence of wearing a fully zipped coat.

A behavior chain has three important characteristics: (a) It involves the performance of a specific series of discrete responses; (b) the performance of each behavior in the sequence produces a stimulus change in the environment in such a way that it yields conditioned reinforcement for the preceding response, and serves as an S^D for the next response; and (c) the responses within the chain must be performed in a specific sequence, and in close temporal succession.

Behavior Chains with a Limited Hold

A **behavior chain with a limited hold** is a chain that must be completed within a specified time to produce reinforcement. Behavior chains with limited holds are characterized by performance that is accurate *and* proficient. An assembly task on a production line exemplifies a chain with a limited hold. Suppose that to meet the production requirements of the job, an employee has to assemble 30 couplers onto 30 shafts within 30 minutes (one per minute). If the coupler is placed on the shaft, the retaining clip applied, and the unit forwarded to the next person in line within the prescribed period, reinforcement is delivered. In a behavior chain with a limited hold, the person must not only have the prerequisite behaviors in his repertoire, but also must emit those behaviors in close temporal succession to obtain reinforcement.

Figure 23.2 Discriminative stimuli, responses, and conditioned reinforcers in a four-response chain.

Discriminative Stimulus	Response	Conditioned Reinforcement
S_1 "Put on your coat."	R_1 Obtain coat from closet	Coat in hands
S_2 Coat in hands	R_2 Place one arm in sleeve	One arm in sleeve
S_3 One arm in sleeve/one arm out	R_3 Place second arm in sleeve	Coat on
S_4 Coat on	R_4 Zip up the coat	Teacher praise

When working with a client who completes a chain in the correct sequence, but fails to meet the limited hold criterion, the behavior analyst should measure the latency and duration of each component response. If this assessment reveals that the client is taking too much time to complete one or more component responses, the behavior analyst can implement contingencies focusing on the proficiency of performance. Setting a time criterion for the onset and completion of each response within the chain is one way to do this; another is to set a time criterion for completion of the total chain.

RATIONALE FOR CHAINING

Whereas *behavior chain* denotes a particular sequence of stimuli and responses ending in reinforcement, **chaining** refers to various methods for linking specific sequences of stimuli and responses to form new performances. In forward chaining, behaviors are linked beginning with the *first* behavior in the sequence. In backward chaining, the behaviors are linked beginning with the *last* behavior in the sequence (Murphy & Lupfer, 2011). Both of these procedures and their variations are discussed in detail later in this chapter.

New behavior chains are taught to people for several reasons. First, an important educational aspect of programs for students with development disabilities—and, for that matter, novice learners—is to increase independent-living skills (e.g., using public utilities, taking care of personal needs, executing travel skills, socializing appropriately). As these skills develop, the learner functions more effectively in least restrictive environments or participates in activities without adult supervision. Complex behaviors that are developed with chaining procedures allow individuals to function more independently.

Second, chaining provides the means by which a series of discrete behaviors can be combined to form a sequence of responses that occasion the delivery of reinforcement. That is, chaining can be used to add behaviors to an existing behavioral repertoire. For example, assume that a person with a developmental disability consistently seeks assistance from a coworker or job coach while completing an assembly task. A chaining procedure could be used to increase the number of tasks in the assembly that must be performed before reinforcement is delivered. To do so, the teacher could give the individual a written or pictorial list of the parts that must be assembled to complete the job. When the first task is finished, the worker crosses off the first word or picture on the list and then proceeds to the next task.

In behavioral terms, the first word or picture on the list serves as the S^D to occasion the response of completing the first task. That response, in the presence of the word or picture on the list, terminates the initial stimulus and produces the next stimulus, the second word or picture on the list. The completion of the second task serves as conditioned reinforcement for completing the task and produces the onset of the third S^D. In this way, the chaining procedure enables simple behaviors to be combined into a longer series of complex responses.

Finally, chaining can be combined with other behavior change tactics to build more intricate and adaptive repertoires. McWilliams, Nietupski, and Hamre-Nietupski (1990) taught three students with moderate intellectual disabilities a 24-step chain of bedmaking skills with an intervention that combined chaining with modeling and error correction. Lambert and colleagues (2016) used a forward chaining procedure that incorporated prompting, instructions, and modeling to teach three complex chains of basketball skills to a middle school student with autism.

ESTABLISHING BEHAVIOR CHAINS WITH TASK ANALYSIS

Before individual behaviors can be linked in a chain, the behavior analyst must construct and validate a task analysis of the components of the behavioral sequence, and assess the learner's mastery level of each behavior in the task analysis. **Task analysis** involves breaking a complex task into smaller, teachable units, the product of which is a series of sequentially ordered steps or tasks.

Constructing and Validating a Task Analysis

The purpose of constructing and validating a task analysis is to determine the sequence of behaviors that are necessary and sufficient to complete a given task efficiently. The sequence of behaviors that one person performs may not be identical to what another person needs to achieve the same outcome. A task analysis should be individualized according to the age, skill level, and prior experience of the learner (Brown, McDonnell, & Snell, 2016). Further, a few task analyses comprise a limited number of major steps, with each major step containing four to five subtasks. Figure 23.3 illustrates a task analysis for making a bed, developed by McWilliams and colleagues (1990).

Researchers and practitioners typically construct and validate task analyses using one method or a combination of four methods: observe a competent performer, execute the task yourself, ask an expert, or trial and error.

Observe a Competent Performer

Test, Spooner, Keul, and Grossi (1990) produced a task analysis for using a public telephone by observing two adults perform the task, and then validating that sequence by training a person with developmental disabilities to use the task analysis. Based on the training, the researchers made subsequent modifications to the original sequence of tasks.

Execute the Task Yourself

A study by van den Pol and colleagues (1981) serves as a good example of how to determine and validate a task analysis yourself. To identify the behaviors required to eat in fast-food restaurants, their research team ate in a variety of restaurants and self-recorded their activities. The researchers next combined their individual findings into a 22-item task analysis that covered differences among restaurants and then organized the steps into four skills clusters: locating, ordering, paying and eating, and exiting (see Figure 30.1). The advantage of self-performing the task is threefold: (1) It provides the opportunity to come into contact with all task demands of the sequence, (2) it generates a clearer idea of the behaviors to be taught and their associated S^Ds, and (3) it refines the description of the response topography necessary for the learner

Figure 23.3 Task analysis of bedmaking skills.

Given an unmade bed that included a messed-up bedspread, blanket, pillow, and flat sheet and a fitted sheet, the student will:

Section 1: bed preparation
1. remove pillow from bed
2. pull bedspread to foot of bed
3. pull blanket to foot of bed
4. pull flat sheet to foot of bed
5. smooth wrinkles of fitted sheet

Section 2: flat sheet
6. pull top of flat sheet to headboard
7. straighten right corner of flat sheet at foot of bed
8. repeat #7 for the left corner
9. make sheet even across top end of mattress
10. smooth wrinkles

Section 3: blanket
11. pull top of blanket to headboard
12. straighten right corner of blanket at foot of bed
13. repeat #12 for the left corner
14. make blanket even with sheet across top end of mattress
15. smooth wrinkles

Section 4: bedspread
16. pull top bedspread to headboard
17. pull right corner of bedspread at foot to floor
18. repeat #17 for left side
19. make bedspread even with floor on both sides
20. smooth wrinkles

Section 5: pillow
21. fold top of spread to within 4 inches of the pillow width
22. place pillow on folded portion
23. cover pillow with the folded portion
24. smooth bedspread wrinkles on and around the pillow

From "Teaching Complex Activities to Students with Moderate Handicaps Through the Forward Chaining of Shorter Total Cycle Response Sequences," by R. McWilliams, J. Nietupski, & S. Hamre-Nietupski, 1990, *Education and Training in Mental Retardation, 25,* p. 296.

to use the sequence most efficiently. Figure 23.4 shows how an initial 7-step sequence for tying a shoe was expanded to 14 steps after self-performing the behavior (cf. Bailey & Wolery, 1992).

Expert Input

Persons or associations with recognized and demonstrated expertise can be a valuable resource to consult when validating a task analysis. For example, Stokes, Luiselli, and Reed (2010) based a 10-step task analysis for football tackling skills on recommendations by the *American Football Coaches Association* (1995). Stokes, Luiselli, Reed, and Fleming (2010) asked five college offensive line coaches to validate a 10-step task analysis for pass-blocking by offensive linemen (see Figure 23.5).

Trial and Error

A systematic trial-and-error procedure can assist the behavior analyst in developing a task analysis. Using a systematic trial-and-error method, an initial task analysis is generated and then refined and revised as it is being tested. With revisions and refinements obtained through field tests, a more functional and appropriate task analysis can be achieved. For instance, as mentioned earlier, Test and colleagues (1990) generated their initial task analysis for using a public telephone by watching two adults complete the task. They subsequently carried the process a step further by asking a person with developmental disabilities to perform the same task and then modified the task analysis.

Regardless of the method used to sequence the steps, the S^Ds and corresponding responses must be identified. However, being able to perform a response is not sufficient; the individual must be able to discriminate the conditions under which a given response should be performed. Listing the discriminative stimuli and associated responses helps the trainer determine whether naturally occurring S^Ds will evoke different or multiple responses.

Assessing Mastery Level

Mastery level is assessed to determine which components of the task analysis the learner can perform independently.

Figure 23.4 Initial and expanded steps for teaching shoe tying.

Shorter sequence[a]	Longer sequence[b]
1. Partially tighten shoe laces.	8. Pinch lace.
2. Pull shoe laces tight—vertical pull.	9. Pull lace.
3. Cross shoe laces.	10. Hang lace ends from corresponding sides of shoe.
4. Tighten laces—horizontal pull.	11. Pick up laces in corresponding hands.
5. Tie laces into a knot.	12. Lift laces above shoe.
6. Make a bow.	13. Cross right lace over the left to form a tepee.
7. Tighten bow.	14. Bring left lace towards student.
	15. Pull left lace through tepee.
	16. Pull laces away from each other.
	17. Bend left lace to form a loop.
	18. Pinch loop with left hand.
	19. Bring right lace over the fingers—around loop.
	20. Push right lace through hole.
	21. Pull loops away from each other.

Sources: (a) Santa Cruz County Office of Education, Behavioral Characteristics Progression. Palo Alto, California, VORT Corporation, 1973. (b) Smith, D. D., Smith, J. O., & Edgar, E. "Research and Application of Instructional Materials Development." In N. G. Haring & L. Brown (Eds.), *Teaching the Severely Handicapped* (Vol. 1). New York: Grune & Stratton, 1976. From *Teaching Infants and Preschoolers with Handicaps*, p. 47, by D. B. Bailey & M. Wolery, 1984, Columbus, OH: Charles E. Merrill. Used by permission.

Figure 23.5 Task analysis for pass-blocking by offensive linemen.

Step		Description
1	Stance	Feet shoulder-width apart, 50% bend at the knees, buttocks down over heels, center-side hand touching the ground
2	Split	Sideline foot back slightly with toe to heel of center-side foot
3	First step	Make play-side bucket step while maintaining balance
4	Helmet contact	Helmet facemask makes contact with opponent at the top of jersey number
5	Hand placement	Pointer fingers at 11:00 o'clock and 1:00 o'clock with thumbs pointed towards opponent's Adam's apple
6	Hand position	Hands maintained inside opponent's shoulder pads and torso frame
7	Arm extension	Arms extended outward beyond 45 degrees
8	Hip follow through	Hips rolled towards opponent's belly button
9	Leg drive	Maintain continuous running leg pump until whistle sounds
10	Hand contact	Hands in contact with opponent's body until whistle sounds

From "Behavioral Coaching to Improve Offensive Line Pass-Blocking Skills of High School Football Athletes," by J. V. Stokes, J. K. Luiselli, D. D. Reed, and R. K. Fleming, 2010, *Journal of Applied Behavior Analysis*, *43*, p. 465. Copyright 2010 by the Society for the Experimental Analysis of Behavior, Inc. Used by permission.

There are two principal ways to assess a person's mastery level of task analysis behaviors prior to training—the single-opportunity method and the multiple-opportunity method.

Single-Opportunity Method

The single-opportunity method is designed to assess a learner's ability to perform each behavior in the task analysis in correct sequence. Lambert and colleagues (2016) used a single-opportunity method to assess target chains related to teaching functional basketball skills to an adolescent with autism. If a behavior in the chain was emitted correctly, confederate players on the court repositioned themselves so as to serve as the S^D for the next behavior. If that behavior was not performed, the assessment probe terminated.

Figure 23.6 is an example of a form used to record a learner's ability to insert a hearing aid. Specifically, a plus sign (+) or minus sign (−) is marked for each behavior that is correctly or incorrectly emitted.

Assessment in this example began when the teacher said, "Tom, put on your hearing aid." Tom's responses to the steps in the task analysis were then recorded. The figure shows Tom's data for the first 4 days of assessment. On Day 1, Tom opened the hearing aid container and removed the harness; each of these steps was performed correctly, independently, sequentially, and within a 6-sec time limit. However, Tom then attempted to put the hearing aid harness over his head (Step 5) without first doing Steps 3 and 4. Because he continued to perform this behavior for more than 10 seconds, the teacher stopped the assessment and scored Steps 3 and 4 and all remaining steps as incorrect. On Day 2, Tom was stopped after Step 1 because he performed a subsequent step out of order. On Days 3 and 4, the assessment was discontinued after Step 4 because Tom took more than 6 seconds to perform Step 5. Given a criterion for mastery of 100% accuracy within 6 seconds over three consecutive probes, the data indicate that Tom only met the criterion for Step 1 (the three +'s for Step 2 were not recorded consecutively).

Multiple-Opportunity Method

The multiple-opportunity method of task analysis assessment evaluates the learner's level of mastery across all the behaviors in the task analysis. If a step is performed incorrectly or out of sequence, or if the time limit for completing the step is exceeded, the behavior analyst completes that step for the learner and then positions her for the next step. Each step performed correctly is scored as a correct response, even if the learner erred on the previous steps.

Figure 23.7 shows an example of the multiple-opportunity method for assessing a student's (Marc's) morning-arrival routine across 17 tasks and 10 days. Brown et al. (2016) validated their task analysis by (a) watching other students complete the morning-arrival routine, (b) conferring with Marc's teacher to

Figure 23.6 Task analysis data sheet for single-opportunity assessment of inserting a hearing aid.

Task Analysis Assessment of Inserting a Hearing Aid
Instructional cue: "Put on your hearing aid"
Teacher: Christine
Assessment method: Single opportunity
Student: Tom

	Date			
Step Behavior	10/1	10/2	10/3	10/4
1. Open container	+	+	+	+
2. Remove harness	+	−	+	+
3. Arm 1/Strap 1	−	−	+	+
4. Arm 1/Strap 2	−	−	+	+
5. Harness over head	−	−	−	−
6. Fasten harness	−	−	−	−
7. Unsnap pocket	−	−	−	−
8. Remove aid from container	−	−	−	−
9. Insert aid into pocket	−	−	−	−
10. Snap pocket	−	−	−	−
11. pick up earmold	−	−	−	−
12. Insert earmold into ear	−	−	−	−
13. Turn on aid	−	−	−	−
14. set the control	−	−	−	−
Percentage of Steps Correct	14%	7%	28%	28%

Materials: Hearing aid container, harness, earmold
Response latency: 6 seconds
Recording key: + (correct) − (incorrect)
Criterion: 100% correct performance for 3 consecutive days

From "Teaching Severely Multihandicapped Students to Put on Their Own Hearing Aids," by D. J. Tucker and G. W. Berry, 1980, *Journal of Applied Behavior Analysis, 13*, p. 69.

Figure 23.7 Task analysis data collection form for multiple-opportunity assessment of student's morning arrival routine.

Teachers: Wharton, Kwan **Instructional cue**: Arrival at school by school bus; bus stops, kids stand up
Student: Marc **Settings**: Bus arrival area, sidewalk, lobby, hallway, classroom **Target**: Morning arrival routine
Day(s): Daily at arrival **Stage of learning**: Acquisition **Teaching method**: Constant time delay (0, 4 seconds)
Probe schedule: First Tuesday of each month **Baseline/Probe method**: Multiple opportunity task analytic assessment (4-sec. latency)

Dates →	9/21	9/22	9/23	9/24	9/27	9/28	9/30	10/1	10/4	10/5
Delayed prompt →			0	0	4	4	4	4	4	
Task Steps ↓										
1. Get off bus.	−	−	√	√	√	√	√	√	+	+
2. Open and walk through door (help OK).	−	−	√	√	√	√	√	√	√	−
3. Walk down the hallway (through lobby to left).	+	−	√	√	√	√	√	√	√	−
4. Open Ms. Kwan's door, go in.	−	−	√	√	√	√	√	√	√	−
5. Wave to Ms. Kwan*.	−	−	√	√	√	√	√	+	+	+
6. Find empty cubby, take off backpack.	−	−	√	√	+	+	+	+	+	+
7. Put backpack inside cubby (on floor).	+	−	√	√	√	+	+	+	+	+
8. Take off jacket.	−	−	√	√	√	√	√	√	+	−
9. Hang it up (empty hook).	−	−	√	√	√	√	+	√	+	−
10. Go to your schedule, get card (first card).	+	+	√	√	√	+	+	+	+	+
11. Go to ______ and get started (first card).	−	−	√	√	√	√	√	√	√	−
12. Go to schedule, get rainbow rug card (when teacher rings bell).	−	−	√	√	√	√	√	√	√	−
13. Go sit on rainbow rug (criss-cross).	−	+	√	√	√	√	+	+	+	+
14. Listen and do _____ (use circle schedule).	−	−	√	√	√	√	√	√	√	√
15. Go to schedule, get Ms. Wharton's room card (when circle done).	−	−	√	√	√	√	+	+	+	+
16. Find Ms. Kwan, say good-bye*.	−	−	√	√	√	√	√	√	+	+
17. Go to Ms. Wharton's room.	−		√	√	√	+	√	+	√	−
Total independent	**3**	**2**	**0**	**0**	**1**	**4**	**6**	**7**	**10**	**8**
Baseline/Teach/Probe	**B**	**B**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**P**

Date	**Teacher**	**Anecdotal Comments**
91/23	JW	Waited for help most steps
9/25	JW	Sleepy, ear infection mods
10/4	LK	He's more sure
10/5	LK	Great probe!
	[Located on back of task analysis]	

Materials: Arrival schedule, backpack, jacket
Latency Period: 0 sec., 4 sec.
Criterion: 10 of 15 steps correct (67%) for 3 of 5 teaching days

Recording Key: Test: + correct, − incorrect; Teach: + unprompted correct,
√ prompted correct (gestural/partial physical prompt). − unprompted/prompted error:
NR no response
* Social enrichment steps

gain her perspective on the tasks, and then (c) "piloting" the task analysis with Marc to ensure that it was individualized. After accomplishing these three validation procedures, it was noted that a couple of tasks were still too difficult for Marc, so each of those steps was further subdivided into two steps. Brown et al. recommend that practitioners developing a task analysis use a systematic process to do so (see Box 23.1).

The key to using the multiple-opportunity method for a task analysis assessment is to ensure that teaching is not commingled with assessment. That is, if the learner is unable to perform a task, the behavior analyst would complete that step, position the learner for the next step, and continue with remaining steps.

Single-opportunity and multiple-opportunity methods can be effective ways to determine mastery of initial skills in a behavior chain. Of the two, the single-opportunity method is the more conservative measure because the assessment terminates at the first step at which performance breaks down. It also provides less information to the teacher once instruction is initiated; but takes less time to conduct, especially if the task analysis is long, and it reduces the likelihood of learning taking place during assessment (Brown et al., 2016). The multiple-opportunity method takes more time to complete, but it provides the behavior analyst with more information. That is, the teacher could learn which steps in the task analysis the learner has already mastered, thus eliminating the need to instruct steps already in the learner's repertoire. In the next section, we address the third component of chaining: teaching the individual to perform each step of the chain in close temporal succession.

BEHAVIOR CHAINING METHODS

After the task analysis has been constructed and validated and the criterion for success and the data collection procedures have been determined, the next step is to decide which chaining procedure to use in teaching the new sequence of behaviors. The behavior analyst has four options: forward chaining, total-task chaining, backward chaining, and backward chaining with leap aheads.

Forward Chaining

In **forward chaining**, the behaviors identified in the task analysis are taught in their naturally occurring order. Specifically, reinforcement is delivered when the predetermined criterion for the performance of the first behavior in the sequence is achieved.

BOX 23.1

How the Task Analysis for Marc's Morning Routine Was Developed

Analyzing a chain of events and breaking these tasks down into teachable steps is not a trivial process. Consider how Marc's teachers developed his task analysis and the data collection form for his morning routine (Figure 23.7). Ms. Wharton, the special education teacher, and the school's autism specialist watched other children perform the morning routine tasks and then analyzed the steps. They identified the relevant stimuli Marc needed to attend to and the responses they wanted Marc to learn. Ms. Wharton also asked Marc's kindergarten teacher, Ms. Kwan, to double-check the preliminary task analysis steps by using them when observing other students arrive in the morning. Ms. Wharton then piloted the task analysis with Marc to be sure the steps made sense for him. When she found that some of the steps were too hard for Marc to do alone (e.g., Step 2: Open the door of the building), she modified those steps by either breaking that step into two smaller steps or adding adult/peer assistance. Marc's mother, who was familiar with the arrival routine, then reviewed the task analysis and added her own ideas. The speech-language pathologist suggested adding a greeting. The 17-step task analysis for Marc's arrival routine reflects the cumulative input of the teachers, the speech therapist, and Marc's mother, thus significantly enhancing the validation process.

Teams will develop more successful task analyses if they follow a systematic process:

1. Select a needed skill by using ecological inventory results to identify a functional and age-appropriate skill that is an important target for a particular student.
2. Define the target skill simply, including a description of the settings and materials most suited to the natural performance of the task.
3. Perform the task and observe peers performing the task, using the chosen materials in the natural settings while noting the steps involved.
4. Adapt the steps to suit the student's disabilities and skill strengths; employ as needed the principle of partial participation, the only-as-specialized-as-necessary rule, and component analysis to design a task analysis that is both age appropriate and functional.
5. Validate the task analysis by having the student perform the task, but provide assistance on steps that are unknown so that performance of all of the steps can be viewed.
6. Revise the task analysis so that it works; explore adding simple, non-stigmatizing adaptations to steps that appear to be unreasonable in an unadapted form.
7. Write the task analysis on a data collection form so that steps (a) are stated in terms of observable behavior; (b) result in a visible change in the product or process; (c) are ordered in a logical sequence; (d) are written in second-person singular so that they could serve as verbal prompts (if used); and (e) use language that is not confusing to the student, with the performance details that are essential to assessing enclosed in parentheses.

Adapted from *Instruction of Students with Severe Disabilities* (8th ed.), pp. 158–159, by F. E. Brown, J. J. McDonnell, and M. E. Snell, 2016.

Thereafter, reinforcement is delivered for criterion completion of Steps 1 and 2. Each succeeding step requires the cumulative performance of all previous steps in the correct order.

For example, when the first step, "Pinch lace," is performed accurately 3 consecutive times, reinforcement is delivered (refer to Figure 23.4). Next, reinforcement would be delivered when that step and the next one, "Pull lace," are performed to the same criterion. Then, "Hang lace ends from corresponding sides of shoe" would be added, and all three steps would have to be performed correctly before reinforcement is delivered. Ultimately, all 14 steps in the task analysis should be performed in a similar manner. However, at any given training step a variety of response prompting and other strategies may be used to occasion the response.

Longer chains of behaviors can be broken down into smaller chains, or skill clusters, each of which can be taught in a manner similar to that used with a single-response unit. When one skill cluster is mastered, it is linked to the next. The final response in the first skill cluster sets the occasion for the first response in the second skill cluster. Essentially, in this variation skill clusters become the analogue for units of behaviors, and these clusters are linked.

McWilliams and colleagues (1990) combined skill clusters in forward chaining to teach bedmaking skills to three students with developmental disabilities. Based on a task analysis of bedmaking responses, five skill clusters were identified, each containing four to five subtasks (see Figure 23.3). Once baseline was collected showing minimal accuracy with completing the entire chain in the presence of the cue, "Show me how to make the bed," an instructional procedure was chosen that involved teaching the clusters one at a time by breaking the complex behaviors into smaller chains (see Figure 23.8).

Initial training began with the teacher demonstrating the chain of tasks for Sections 1 and 2. Subsequent training involved teaching the prior section, the current section targeted for instruction, and the next section in the sequence. For example, if training was to be conducted on Section 2, Sections 1, 2, and 3 were demonstrated. Next, the students practiced the targeted sequences 2 to 5 times. When the sequence was performed correctly, praise was delivered. When an error occurred, a three-part correction procedure involving verbal redirection, redirection plus modeling, and/or physical guidance was initiated until the trial ended on a correct response. When the first skill cluster was mastered (S_1), the second skill cluster (S_2) was introduced, followed by the third (S_3), the fourth (S_4), and so on.

The results indicated that the forward chaining procedure was effective in teaching the students bedmaking skills. That is, all the students were able to make their beds independently, or with only minimal assistance, when the teacher directed them to do so. Furthermore, all of them were able to make their beds in the generalization setting (i.e., the home).

McWilliams and colleagues (1990) and Guercio and Cormier (2015) illustrate three main advantages of forward chaining: (a) It can be used to link smaller chains into larger ones; (b) it is easy, so teacher implementation in the classroom is enhanced; and (c) it can be combined with other behavior change procedures (e.g., fading).

Despite its advantages, forward chaining has potential limitations. For example, it is possible that the analyst, after noting the learner has performed the targeted training steps, may

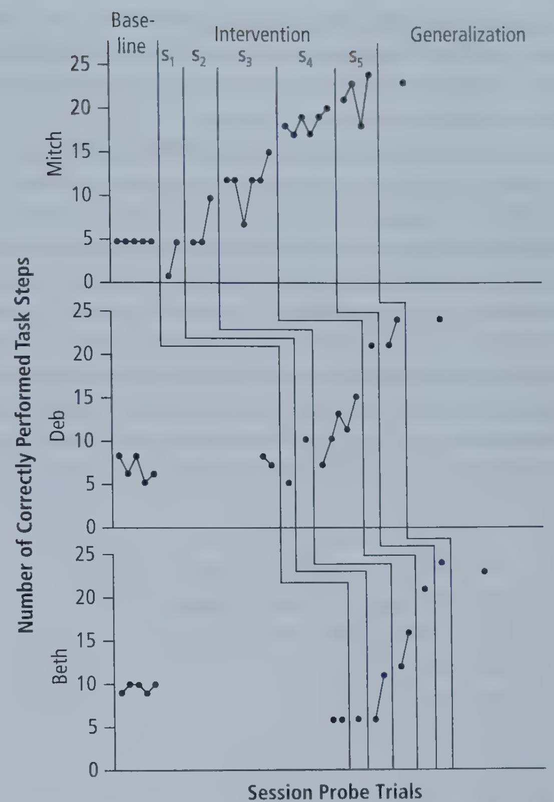


Figure 23.8 Number of correctly performed task steps across baseline, intervention, and generalization probe trials.

From "Teaching Complex Activities to Students with Moderate Handicaps Through the Forward Chaining of Shorter Total Cycle Response Sequences," by R. McWilliams, J. Nietupski, & S. Hamre-Nietupski, 1990, *Education and Training in Mental Retardation*, *25*, p. 296. Copyright 1990 by the Council for Exceptional Children. Reprinted by permission.

eliminate opportunities for additional practice of these steps, thus risking extinction of the chain due to lack of reinforcement or the weakening of an S^D. Conversely, the analyst may retain the learner in forward chaining intervention longer than necessary, thereby hampering progress on additional skills to be learned. To mitigate either or both of these limitations, Lambert et al. (2016) suggest that analysts probe the entire chain periodically to assess mastery and retention.

Our discussion thus far has focused on teaching behavior chains through direct teacher instruction. However, evidence suggests that chained responses can also be learned through observation (Wolery, Ault, Gast, Doyle, & Griffen, 1991). Griffen, Wolery, and Schuster (1992) used a constant time delay (CTD) procedure to teach students with intellectual disabilities a series of chained tasks in food preparation behaviors. During CTD with one student to whom the chained response was taught, two other students watched. Results showed that the two observing students learned at least 85% of the correct steps in the chain even though direct instruction on those steps had not occurred.

Total-Task Chaining

Total-task chaining (sometimes called *total-task presentation* or *whole-task presentation*) is a variation of forward chaining in which the learner receives training on each step in the task analysis during

every session. The trainer provides assistance for any steps the learner is unable to perform independently, and the chain is trained until the learner is able to perform all the behaviors in the sequence to the predetermined criterion. Depending on the complexity of the chain, the learner's repertoire, and available resources, physical assistance and/or graduated guidance may be incorporated.

Werts, Caldwell, and Wolery (1996) used peer models and total-task chaining to teach skills such as operating an audiotape, sharpening a pencil, and using a calculator to three elementary-level students with disabilities who were enrolled in general education classrooms. Response chains were individualized for each student based on teacher-recommended sequences that the students had to learn. Each session was divided into three parts: (a) The students with disabilities were probed on their ability to perform the total-task response chain; (b) a competent peer model demonstrated the chain in its entirety while simultaneously describing each step; and (c) the student partner was probed again to determine performance of the chain. During the probes before and after peer modeling, the students with disabilities were directed to complete the chain. If a student was successful, a correct response was noted, but no feedback was delivered. If a student was unsuccessful, the student's view was blocked temporarily while the teacher completed that step in the chain, and then the student was redirected to complete the remaining steps. Each behavior in the response chain was scored.

All three students learned to complete the response chain after peer modeling, and reached criterion on the response chain (100% correct for 2 out of 3 days) over the course of the study. The results for Charlie, one of the three students in the study, are shown in Figure 23.9.

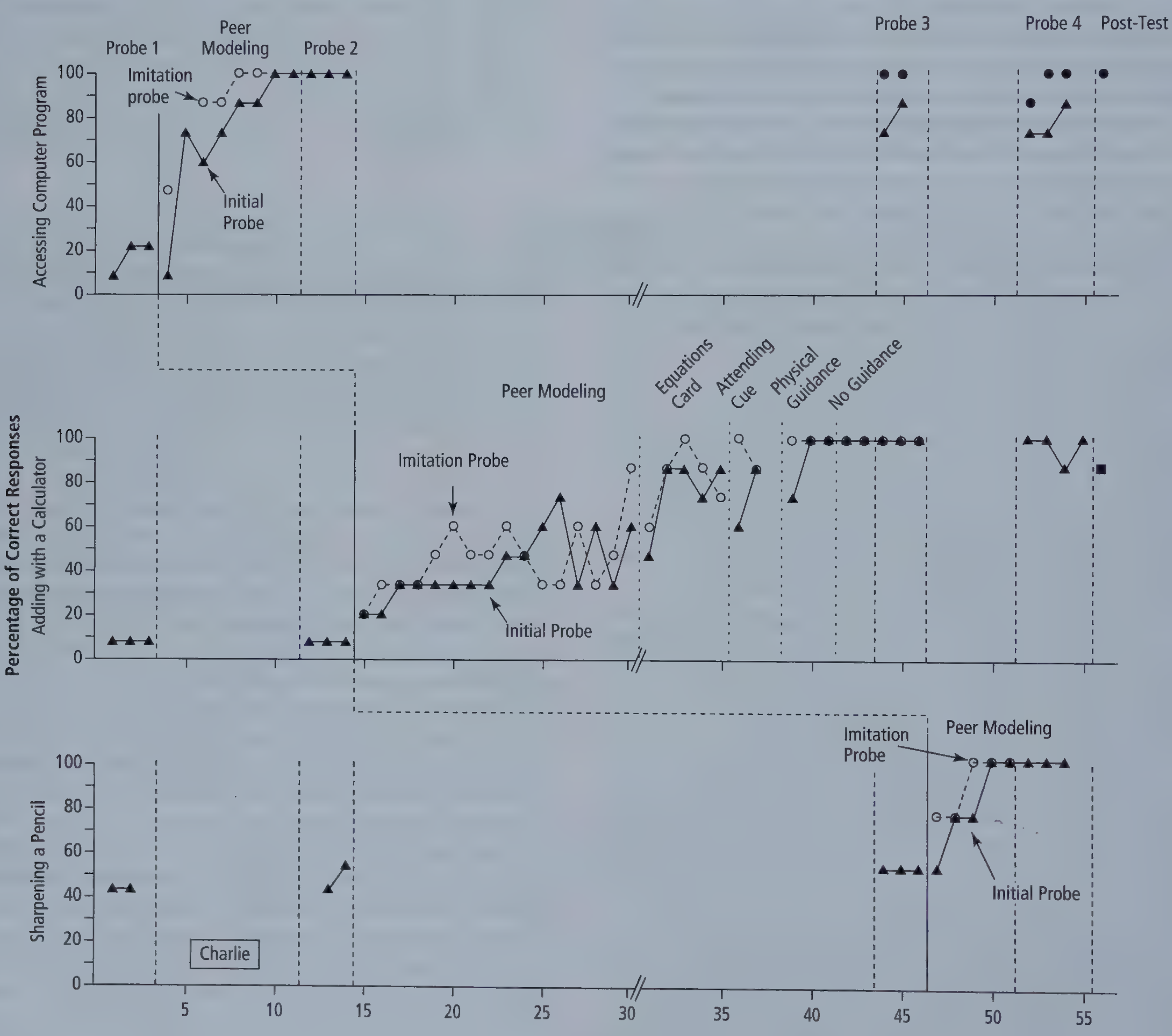


Figure 23.9 Percentage of steps performed correctly by Charlie for the three response chains. Triangles represent percentage of steps correct in initial probes; open circles represent percentage of steps correct in imitation probes.

From "Peer Modeling of Response Chains: Observational Learning by Students with Disabilities," by M. G. Werts, N. K. Caldwell, and M. Wolery, 1996, *Journal of Applied Behavior Analysis, 29*, p. 60. Copyright 1996 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

Test and colleagues (1990) used total-task chaining with a least-to-most prompting procedure to teach two adolescents with severe intellectual disabilities to use a public telephone.[1] After identifying and validating a 17-step task analysis, baseline data were obtained by giving the students a 3 × 5 index card with their home phone number and two dimes, and directing them to "phone home." During training, a least-to-most prompting procedure consisting of three levels of prompts (verbal, verbal plus gesture, and verbal plus guidance) was implemented when errors occurred on any of the 17 steps that made up the task analysis. Each instructional session consisted of two training trials followed by a probe to measure the number of independent steps completed. In addition, generalization probes were conducted in two other settings at least once per week.

Results showed that the combination of total-task chaining plus prompts increased the number of steps in the task analysis performed correctly by each student and that the skills generalized to two community-based settings (see Figure 23.10). Test and colleagues (1990) concluded that total-task chaining, especially as used within the context of a two-per-week training regimen, offered benefits to practitioners working with learners in community-based settings.

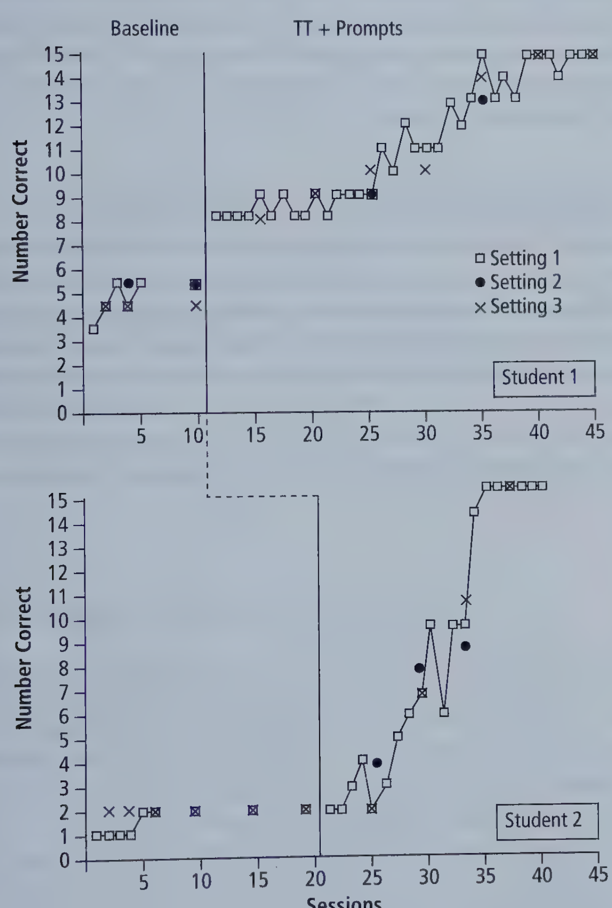


Figure 23.10 Number of correctly performed steps in the public-telephone task analysis by two students during baseline and total-task chaining (TT) plus response prompts. Data points for settings 2 and 3 show performance on generalization probes.

From "Teaching Adolescents with Severe Disabilities to Use the Public Telephone," by D. W. Test, F. Spooner, P. K. Keul, and T. A. Grossi, 1990, *Behavior Modification, 14*, p. 165. Copyright 1990 by Sage Publications. Reprinted by permission.

Backward Chaining

When a **backward chaining** procedure is used, all the behaviors identified in the task analysis are initially completed by the trainer, except for the final behavior in the chain. When the learner performs the final behavior in the sequence at the predetermined criterion level, reinforcement is delivered. Next, reinforcement is delivered when the last and the next-to-last behaviors in the sequence are performed to criterion. Subsequently, reinforcement is delivered when the last three behaviors are performed to criterion. This sequence proceeds backward through the chain until all the steps in the task analysis have been introduced in reverse order and practiced cumulatively.

Pierrel and Sherman (1963) conducted a classic demonstration of backward chaining at Brown University with a white rat named Barnabus. Pierrel and Sherman taught Barnabus to climb a spiral staircase, push down and cross a drawbridge, climb a ladder, pull a toy car by a chain, enter the car and pedal through a tunnel, climb a flight of stairs, run through an enclosed tube, enter an elevator, raise a miniature replica of the Brown University flag, exit the elevator, and finally press a bar for which he received a pellet of food. Barnabus became so famous for his performance of this elaborate sequence that he acquired the reputation of being "the rat with a college education." This chain of 11 responses was trained by initially conditioning the last response in the sequence (the bar press) in the presence of a buzzer sound, which was established as the S^D for the bar press. Then the next-to-last response in the sequence (exiting the elevator) was conditioned in the presence of the elevator at the bottom of the shaft. Each response in the chain was added in turn so that a discrete stimulus served as the S^D for the next response and as the conditioned reinforcer for the preceding response.[2]

To illustrate backward chaining with a classroom example, let us suppose that a preschool teacher wants to teach a student to tie his shoes. First, the teacher conducts a task analysis of shoe tying and puts the component behaviors in logical sequence.

1. Cross over laces on the shoe.
2. Tie a knot.
3. Make a loop with the lace on the right side of the shoe, and hold it in the right hand.
4. With the left hand, wrap the other lace around the loop.
5. Use the index or middle finger of the left hand to push the left lace through the opening between the laces.
6. Grasp both loops, one in each hand.
7. Draw the loops snug.

The teacher begins by training with the last step in the sequence, Step 7, until the student is able to complete it without mistakes for three consecutive trials. After each correct trial at Step 7, reinforcement is delivered. The teacher then introduces the next-to-last step, Step 6, and begins training the student to perform that step and the final step in the chain. Reinforcement is

contingent on the successful performance of Steps 6 and 7. The teacher then introduces Step 5, making sure that the training step and all previously learned steps (i.e., Steps 5, 6, and 7) are executed in the proper sequence prior to reinforcement. The teacher can use supplemental response prompts to evoke correct responding on behavior at any step. However, any supplemental response prompts—verbal, pictorial, demonstration, or physical—introduced during training must be faded later in the program so that the student's behavior comes under stimulus control of the naturally occurring S^Ds.

With backward chaining, the first behavior the learner performs independently produces the terminal reinforcement: The shoe is tied. The next-to-last response produces the onset of a stimulus condition that reinforces that step and serves as an S^D for the last behavior, which is now being established in the learner's behavioral repertoire. This reinforcing sequence is repeated for the remainder of the steps.

Hagopian, Farrell, and Amari (1996) combined backward chaining with fading to reduce the life-threatening behavior of Josh, a 12-year-old boy with autism and intellectual disabilities. Because of medical complications associated with a variety of gastrointestinal factors, frequent vomiting, and constipation during the 6 months prior to the study, Josh refused to ingest any food or liquids by mouth. Indeed, when he did ingest food by mouth previously, he expelled it.

Over the 70-day course of Josh's treatment program, data were collected on liquid acceptance, expulsions, swallows, and avoidance of the target behavior: drinking water from a cup. After baseline was collected on Josh's ability to swallow 10 cc of water, a swallow condition without water was implemented. An empty syringe was placed and depressed in his mouth, and he was directed to swallow. Next, the syringe was filled with a small volume of water, and reinforcement delivered when Josh swallowed the water. In subsequent phases, the volume of water in the syringe was gradually increased from 0.2 cc to 0.5 cc to 1 cc, and ultimately to 3 cc. To earn reinforcement in subsequent conditions, Josh had to emit the target chain using at first 3 cc of water from a cup to ultimately 30 cc of water. At the end of the study, a successful generalization probe was implemented with 90 cc of a mixture of water and juice.

The results showed that the chaining procedure improved Josh's ability to emit the target chain (see Figure 23.11). Hagopian and colleagues explained their procedure this way:

> We began by targeting a preexisting response (swallowing), which was the third and final response in the chain of behaviors that constitute drinking from a cup. Next, reinforcement was delivered for the last two responses in the chain (accepting and swallowing). Finally, reinforcement was delivered only when all three responses in the chain occurred. (p. 575)

A primary advantage of backward chaining is that the learner contacts the chain's terminal reinforcer on every instructional trial. As a direct outcome of reinforcement, the stimulus that is present at the time of reinforcement increases its discriminative properties. Further, the repeated reinforcement of all behaviors in the chain increases the discriminative capability of all the stimuli associated with these behaviors and with the reinforcer. The main disadvantage of backward chaining is that the potential passive participation of the learner in earlier steps in the chain may limit the total number of responses made during any given training session.

Backward Chaining with Leap Aheads

Spooner, Spooner, and Ulicny (1986) described a variation of backward chaining they called backward chaining with leap aheads. **Backward chaining with leap aheads** follows essentially the same procedures as backward chaining, except that not every step in the task analysis is trained. Selected steps are simply probed. The purpose of the leap-aheads modification is to decrease the total training time needed to learn the chain. With conventional backward chaining, the stepwise

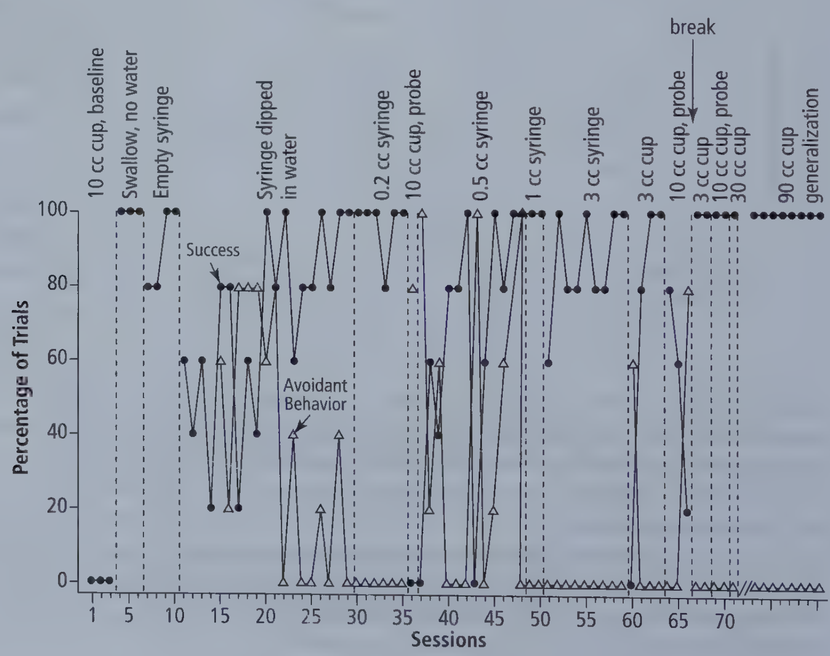


Figure 23.11 Percentage of trials with successful drinking and trials with avoidance behaviors.

From "Treating Total Liquid Refusal with Backward Chaining and Fading," by L. P. Hagopian, D. A. Farrell, and A. Amari, 1996, *Journal of Applied Behavior Analysis*, 29, p. 575. Copyright 1996 by the Society for the Experimental Analysis of Behavior, Inc. Used with permission.

Figure 23.12 Essential features of forward, total-task, backward, and backward chaining with leap aheads.

Forward Chaining	Total-Task Chaining	Backward Chaining	Backward Chaining with Leap Aheads
Behaviors identified in the task analysis are taught in their naturally occurring order. Specifically, reinforcement is delivered when the predetermined criterion for the performance of the first behavior in the sequence is achieved. Trainer completes the remaining tasks in the chain, or assists the learner in doing so. Next, reinforcement is delivered for criterion completion of Steps 1 and 2. Succeeding steps require the cumulative performance of all previous steps in the correct order.	A variation of forward chaining in which the learner receives training on each step in the task analysis during every session. The trainer provides assistance for any steps the learner is unable to perform independently, and the chain is trained until the learner is able to perform all the behaviors in the sequence to the predetermined criterion.	Tasks are taught in a "backward" (last-to-first) sequence. All the task-analysis-identified behaviors are initially completed by the trainer, except for the final behavior in the chain. When the learner performs the final behavior in the sequence at the predetermined criterion level, reinforcement is delivered. Next, reinforcement is delivered when the combined next-to-last and last behaviors in the sequence are performed to criterion. Subsequently, reinforcement is delivered when the combined sequence of the last three behaviors are performed to criterion. This series proceeds in reverse fashion through the chain until all the steps in the task analysis have been introduced in reverse order and practiced cumulatively.	This type of chaining follows essentially the same procedures as backward chaining, except that not every step in the reverse execution of the task analysis is trained. Selected steps are simply probed. Tasks that the learner can perform are skipped, and the trainer "leaps ahead" to the next unmastered task, thereby decreasing total training time for the chain.

repetition of the behaviors in the sequence may slow down the learning process, especially when the learner has mastered a few of the steps in the chain. For instance, in the previous shoe-tying illustration, the child might perform Step 7, the last behavior in the sequence, and then leap ahead to Step 4 because Steps 5 and 6 are in her repertoire. It is important to remember, however, that the learner must still perform Steps 5 and 6 correctly and in sequence with the other steps to receive reinforcement.

Figure 23.12 describes the essential features of forward, total-task, backward, and backward chaining with leap aheads.

CHOOSING A CHAINING METHOD

Forward chaining, total-task chaining, backward chaining, and backward chaining with leap aheads have each shown effectiveness with a wide range of self-care, vocational, recreational, and independent-living behaviors. Which chaining procedure should be the method of first choice? Research conducted to date does not suggest a definitive answer. Furthermore, research does not suggest a clear protocol for whether unfinished links in a chain should be completed by the teacher or the student, or left unfinished until the next training trial. Bancroft, Weiss, Libby, and Ahearn (2011) showed that whereas the majority of students in their study learned a chain faster under the student condition, this result was not universal for all students.

After examining evidence reported between 1980 and 2001, Kazdin (2001) concluded: "Direct comparisons have not established that one [forward chaining, backward chaining, or total-task chaining] is consistently more effective than the other" (p. 49). A study by Slocum and Tiger (2011) found forward and backward chaining procedures to be similarly effective.

Seaver and Bourret (2014) stated, "When all teaching procedures are equally efficient for a particular individual, a practitioner may take student preference . . . teacher preference, and procedural integrity into consideration" (p. 791). Anecdotal evidence, logical analysis, and the behavior analyst's experience with the learner may also provide direction. Miltenberger (2001) and Test et al. (1990) concluded that using total-task chaining may be warranted (a) when the student can perform several of the tasks in the chain, but needs to learn them in sequence; (b) when the student has an imitative repertoire; (c) when the student has moderate to severe disabilities; and/or (d) when the task sequence or cycle is not very long or complex. Likewise, Bancroft et al. (2011) advised considering other factors—the learner's ability to tolerate physical guidance, the attention span of the learner, or the total amount of training time available to devote to learning the chain—when deciding which method to use.

Uncertainty about which approach to follow may be minimized by conducting a personalized task analysis for the learner, systematically applying the single- or multiple-opportunity method to determine the starting point for instruction, and relying on empirically sound data-based studies in the literature. Further, doubt can be reduced by ascertaining if an individual's past history with any of the chaining procedures or variations favors one method over another, and/or by collecting formative evaluation data (i.e., measurements of the target behavior during treatment) and summative evaluation data (i.e., measurement of behavior after treatment) on the proposed approach, to determine its efficacy for that person.

DISRUPTING AND BREAKING BEHAVIOR CHAINS

Thus far, our discussion has focused on procedures for building behavior chains—that is, constructing a sequence of stimulus conditions and associated responses such that when individual components are joined, the chain of behavior results that produces reinforcement. As stated earlier, each stimulus linking two responses in a chain serves the dual functions of conditioned reinforcer for the response that produced it and discriminative stimulus for the next response in the chain.

Practitioners have used chaining to expand the repertoires of a wide range of learners. Knowing how to link behaviors into a practical chain, however, is not always sufficient. Behavior analysts should also know how to disrupt or break behavior chains that are maladaptive or inappropriate (e.g., consuming too much food, texting while driving). Behavior chains can be disrupted or broken with six methods: extinction, satiation, unchaining, interrupting the chain, substituting the S^D, and extending the chain with time delays.

In a series of three experiments, Kuhn, Lerman, Vorndran, and Addison (2006) evaluated the effects of extinction, satiation, and unchaining on responding by children with disabilities in a previously taught two-response chain. The chain in each experiment went like this: S1 (therapist shows child a small box containing a preferred food item and closes the box) → R1 (child signs "open") → S2 (therapist opens box of food) → R2 (child signs "eat" or "popcorn") → (child receives food).

Extinction

To break or interrupt a chain using extinction, the practitioner withholds reinforcement for one or more responses in the chain. Kuhn et al. (2006) applied extinction to the final response in the chain: "R1 ('open') resulted in the box opening as before, but the box was closed and no food was delivered contingent on R2 ('eat' or 'popcorn')" (p. 273). Extinction for R2 resulted in immediate decreases in both R1 and R2 by all four participants (Figure 23.13 shows results for two children). Extending their findings to a two-response chain of aggressive behavior consisting of arm grabbing and hair pulling, Kuhn and colleagues suggested that treating hair pulling with extinction would likely result in decreased arm grabbling.

Satiation

To break or interrupt a chain using satiation, the practitioner provides unlimited access to the reinforcer (an abolishing operation) for the last response in the chain. As a result, the chain's final response will likely abate because of the decreased effectiveness of the reinforcer. Responses earlier in the chain would be expected to decrease as well because the effectiveness of the stimulus changes as conditioned reinforcers for earlier responses depends upon their association with the chain's terminal reinforcer.

Kuhn and colleagues (2006) evaluated this satiation condition. Prior to each session, the child was given free access to food for the duration of "satiation periods" ranging from 25 to 30 minutes. Presession access to the chain's terminal reinforcer decreased both responses in the chain (see Figure 23.14). R2 ("eat") decreased as a direct result of presession access to food; R1 ("open") decreased because presession access to food weakened an opened food box as a conditioned reinforcer.

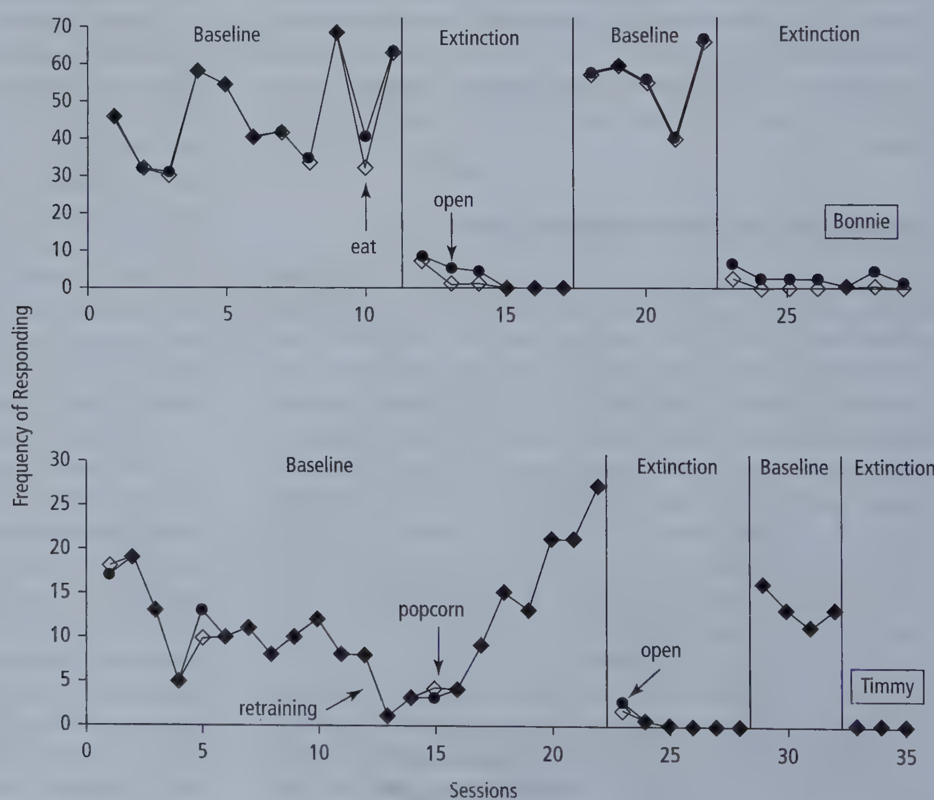


Figure 23.13 Frequency of appropriate "open" and "eat" ("popcorn" instead of "eat" for Timmy) responses during the extinction evaluation for Bonnie (top) and Timmy (bottom).

From "Analysis of Factors That Affect Responding in a Two-Response Chain in Children with Developmental Disabilities," by S. A. C. Kuhn, D. C. Lerman, C. M. Vorndran, and L. Addison, 2006, *Journal of Applied Behavior Analysis, 39*, p. 273. Copyright 2006 by the Society for the Experimental Analysis of Behavior, Inc. Used with permission.

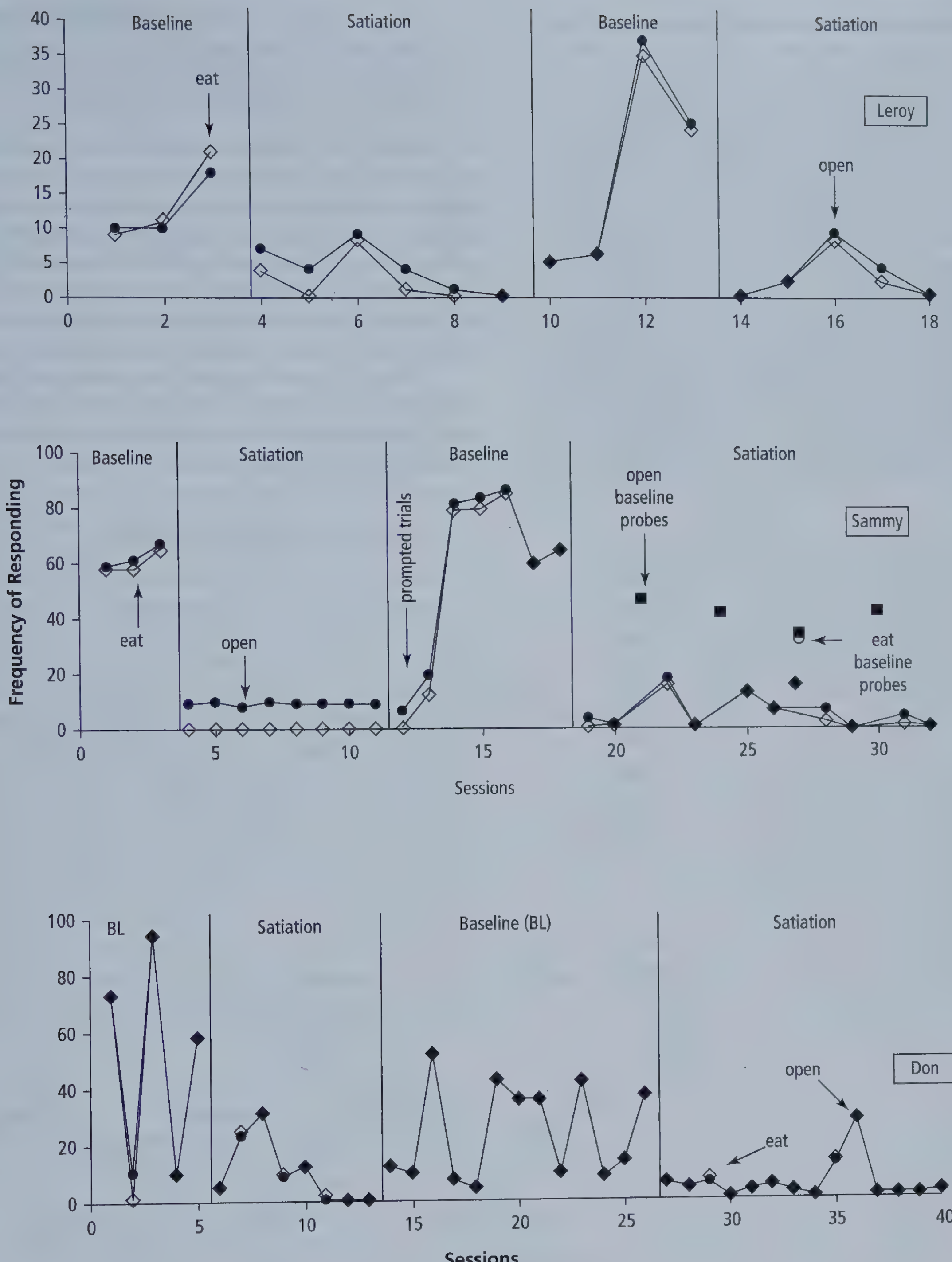


Figure 23.14 Frequency of "open" and "eat" responses during the satiation evaluation for Leroy (top), Sammy (middle), and Don (bottom).

From "Analysis of Factors That Affect Responding in a Two-Response Chain in Children with Developmental Disabilities," by S. A. C. Kuhn, D. C. Lerman, C. M. Vorndran, and L. Addison, 2006, *Journal of Applied Behavior Analysis, 39*, p. 272. Copyright 2006 by the Society for the Experimental Analysis of Behavior, Inc. Used with permission.

Unchaining

Unchaining—a term introduced by Kuhn et al. (2006) for a procedure described by Michael (2000)—occurs when, in the case of a two-response chain, R2 produces reinforcement not only in the presence of S2 (as usual) but also in the absence of S2. In Michael's example, a food-deprived rat's first response (R1) turns on a tone, in the presence of which a second response (R2) produces food. In the unchaining procedure (Michael called it *unpairing*), R1 continues to produce the tone, but R2 results in food regardless of the tone's presence. In theory, R1 will decrease because its immediate consequence, the tone, loses its reinforcing effectiveness as a conditioned reinforcer because it no longer functions as a discriminative stimulus (S^D) for food. Thus, unchaining disrupts R1 while leaving R2 intact.

The unchaining procedure used by Kuhn and colleagues (2006) was implemented as follows:

> Reinforcement was delivered contingent on R2 regardless of when it occurred. That is, the child received a small piece of food contingent on completing the chain (i.e., R1, then R2) and contingent on R2 alone. Thus, R2 produced food regardless of whether the box was open. The purpose of this condition was to look at the effects of unchaining R1 (i.e., communicating to open the box) and R2 (i.e., communicating for food) on the response chain. Initially for all children, the reinforcer for signing "eat" (or "popcorn" for Timmy) outside the chain "open-eat" was placed behind the therapist's back, as well as in the box, so that the therapist could deliver the reinforcer without opening the box. Beginning with Session 16 (first unchaining phase) and Session 42 (second unchaining phase), the reinforcer was placed on a plate on the table in front of Don, next to the box containing additional reinforcers. The reinforcer on the plate was provided contingent on occurrences of the "eat" response outside the context of the chain. (p. 274)

Figure 23.15 shows results for the three children who participated in the experiment. Unchaining had little effect on Bonnie, as she continued to emit both responses in the chain throughout the unchaining condition. In contrast, after three or four sessions of unchaining, Timmy no longer emitted R1. Unchaining had no effect on Don's behavior until researchers modified the procedure by making the food visible on a plate. This resulted in a decrease in R1 ("open") responses.

Although unchaining may show promise as a method for disrupting earlier responses in a chain, Kuhn et al.'s (2006) data do not show that the effects of unchaining are as powerful, or as immediate, as extinction or satiation methods.

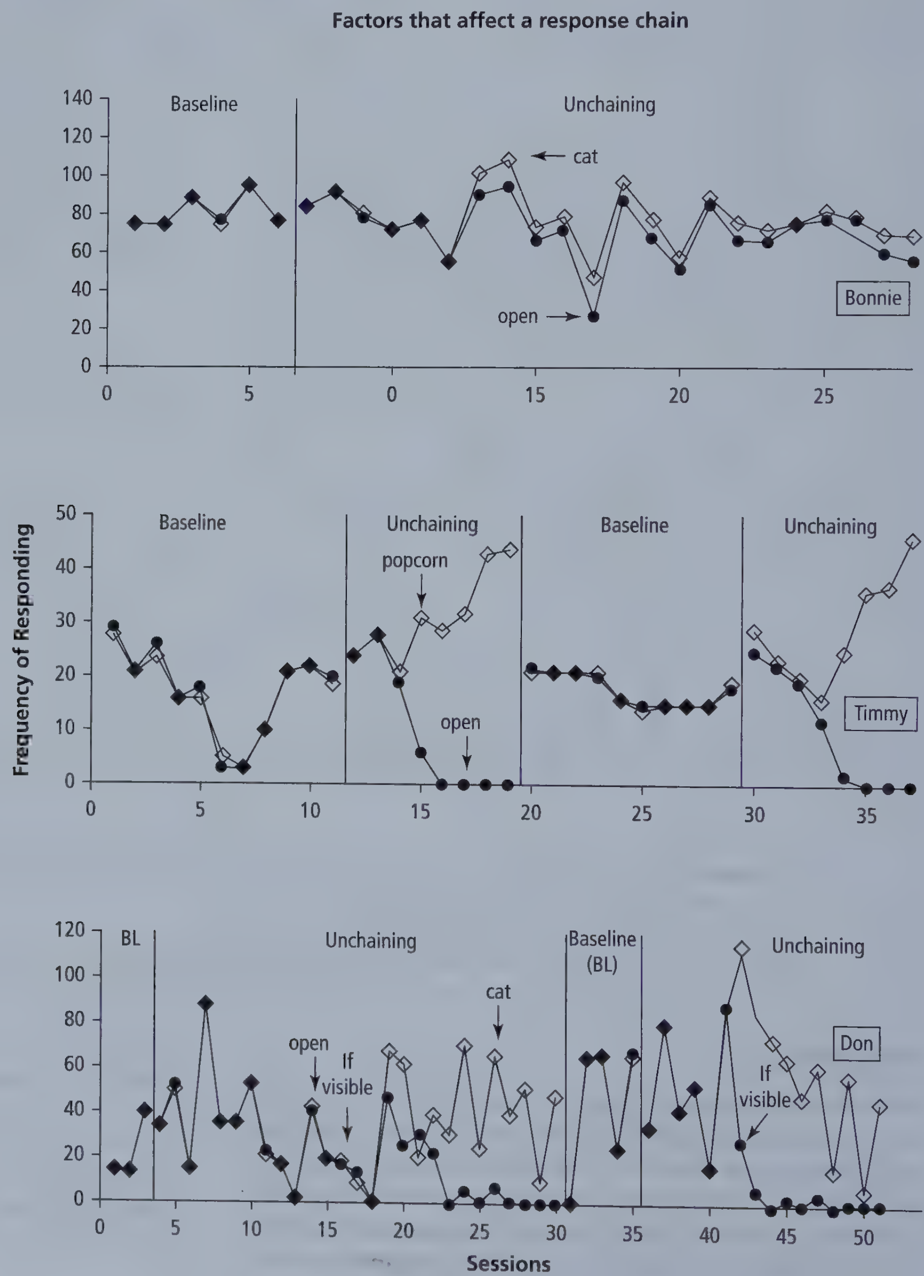


Figure 23.15 Frequency of appropriate "open" and "eat" ("popcorn" for Timmy) responses during the unchaining evaluation for Bonnie (top), Timmy (middle), and Don (bottom).

From "Analysis of Factors That Affect Responding in a Two-Response Chain in Children with Developmental Disabilities," by S. A. C. Kuhn, D. C. Lerman, C. M. Vorndran, and L. Addison, 2006, *Journal of Applied Behavior Analysis*, *39*, p. 275. Copyright 2006 by the Society for the Experimental Analysis of Behavior, Inc. Used with permission.

Interrupting the Chain

The **behavior chain interruption strategy** (BCIS) relies on the participant's skill to initially perform all the critical elements of a chain independently, but the chain is then interrupted, or a link in the chain is made unavailable, at a predetermined link so that another behavior can be emitted or prompted. Initially developed to increase speech and oral responding (Goetz, Gee, & Sailer, 1985; Hunt & Goetz, 1988), the BCIS has been extended into pictorial communication systems (Roberts-Pennell & Sigafoos, 1999), manual signing (Romer, Cullinan, & Schoenberg, 1994), switch activation (Gee, Graham, Goetz, Oshima, & Yoshioka, 1991), and mand training (Albert, Carbone, Murray, Hagerty, & Sweeney-Kerwin, 2012; Hall & Sundberg, 1987).

The BCIS works as follows. First, an assessment is conducted to determine if the person can independently complete a chain of two or more components. Figure 23.16 shows an example of how a toast-making chain, divided into five steps, was assessed across degrees of distress and attempts to complete when specific steps were blocked by an observer. Degree of distress was ranked on a 3-point scale, and attempts were registered using a dichotomous scale (i.e., yes or no).

With respect to using BCIS in applied settings to increase behavior, a chain is selected based on the assumption that when it is interrupted, it will not be unduly distressful for the individual. In other words, care is taken that the interruption does not precipitate the person becoming episodic or self-injurious. After collecting baseline on a target behavior (e.g., vocalizations), the person is directed to start the chain (e.g., "Make toast."). At a predetermined step in the chain—say, Step 3 (e.g., "Push down")—the individual's ability to complete that step is restricted. For example, the practitioner might passively block access to the toaster momentarily. The person would then be prompted, "What do you want?" A vocalization, previously not in the learner's repertoire, would be required for the chain to be completed. That is, the person would be prompted to say, "Push down."

A procedural variation of BCIS, the **interrupted chain procedure** (ICP), entails arranging the environment such that the learner is unable to continue the chain at a predetermined step. A practitioner implementing the ICP in the toast-making chain might conceal the toaster and, after the learner has obtained a slice of bread, ask, "What do you want?" The verbal prompt may function as a motivating operation that evokes the learner's response, "Toaster, please," to obtain the toaster and complete the chain.

Although the exact behavioral mechanisms responsible for improved performance using BCIS and ICP are still subject to scientific inquiry, analysis, and explanation, field-based research efforts, including those examining generalized outcomes (Grunsell & Carter, 2002), have established them as efficacious approaches, especially for persons with severe disabilities. For example, Carter and Grunsell's (2001) review of the literature on the efficacy of BCIS shows that it is positive and beneficial. In their view,

> The BCIS may be seen as empirically supported and complementary to other naturalistic techniques. . . . A small, but growing body of research literature has demonstrated that individuals with severe disabilities have acquired requesting with the BCIS. In addition, research demonstrates that implementation of the BCIS may increase the rate of requesting. (p. 48)

The assumption underlying BCIS assessment is that "persistence in task completion and emotional response to the interruption would serve as operational definitions of high motivation for task completion" (Goetz et al., 1985, p. 23). Because momentarily interrupting a chain detains the learner from obtaining the reinforcer for completing the chain, the procedure may function as a conditioned motivation operation that evokes additional behavior required to continue and complete the chain (see Chapter 16). "By interrupting the chains at the point at which these additional objects (or actions) are needed,

Chain Pretesting

Student's Name: Chuck

Chain: making toast

Date: 5/2

*interrupt at start of step

Sequence of steps	Degree of distress			Attempts to complete		Comments
1. Pull bread from bag*	1 (low)	2	3 (high)	yes	no	
2. Put in toaster	1	2	3	yes	no	
3. Push down*	1	2	3	yes	no	tried to touch knob
4. Take toast out	1	2	3	yes	no	
5. Put on plate*	1	2	3	yes	no	very upset; self-abusive

$\bar{x}$ distress: $\frac{7}{3} = 2.3$ Total 2/3 = 66%

Figure 23.16 Sample score sheet used to evaluate chains.

From "Using a Chain Interruption Strategy to Teach Communication Skills to Students with Severe Disabilities," by L. Goetz, K. Gee, & W. Sailor, *Journal of the Association of Persons with Severe Handicaps, 1985, 13*(1), p. 24. Copyright by the Association for Persons with Severe Handicaps. Used by permission.

practitioners can momentarily establish them as reinforcers and contrive opportunities to teach a wider variety of mands" (Albert et al., 2012, p.73).

In any case, the literature is becoming increasingly clear that the BCIS/ICP consists of several components and features that make it useful as a behavior change tactic in natural settings. Participants exposed to interrupted chain procedures can learn to perform nontrained tasks and responses, thereby increasing the efficiency of the procedure (Albert et al., 2012; Carter & Grunsell, 2001). As Albert el al. state, "By interrupting the chains at the point at which these additional objects (or actions) are needed, practitioners can momentarily establish them as reinforcers and contrive opportunities to teach a wider variety of mands" (p. 73). Figure 23.17 reviews key components and features of BCIS/ICPs.

Substituting an Initial S^D for an Alternative Chain

When a specific S^D sets an inappropriate chain in motion, blocking that S^D completely and substituting the initial S^D for a preferred chain may be effective. For example, if an ashtray next to a favorite chair (S_1) evokes a chain of finding a pack of cigarettes (R_1), finding a lighter (R_2), and firing up a smoke (R_3), replacing the ash tray with a TV remote, a newspaper, or a book may occasion one or more desirable sequences of responses.

Extending the Chain with Time Delays

Martin and Pear (2016) suggested that a behavior chain that reinforces excessive food consumption can be broken by introducing links requiring the individual to place the eating utensil on the table between bites, or introducing a 3- to 5-sec time delay before the next bite can be initiated. In an undesirable chain, "the person gets ready to consume the next bite before even finishing the present one" (p. 111). A more desirable chain separates these components and introduces brief delays.

TROUBLESHOOTING CHAINS

With respect to separating chain components, let us consider the case of a restaurant trainee who has been taught to bus tables. Upon completion of the training program, the trainee was able to perform each of the necessary behaviors in the table-busing chain accurately and proficiently (see Figure 23.18, upper panel). However, on the job site, the new employee started to emit an inappropriate sequence of

Figure 23.17 Key components and features of the BCIS.

- Instruction begins in the middle of the chain sequence, not the start, differing from forward and backward chaining procedures. Because the instruction begins in the middle of the sequence, interrupting the sequence might function as a transitive conditioned motivating operation or negative reinforcer, the removal of which increases behavior.
- Procedurally, the BCIS is predicated on an assessment that verifies that the person is able to complete the chain independently, but experiences moderate distress when the chain is interrupted in the middle of the sequence.
- Verbal prompts are used at the point of interruption (e.g., "What do you want?"), but a full range of response prompts—modeling and physical guidance—has also been employed.
- Interruption training occurs in the natural setting (e.g., a water basin for washing hair; the microwave oven for making cookies).
- Maintenance, generalization, and social validity data, although not overwhelmingly evident in every study, are sufficiently robust to suggest that the BCIS be included with other interventions (e.g., mand-model, time delay, and incidental teaching).

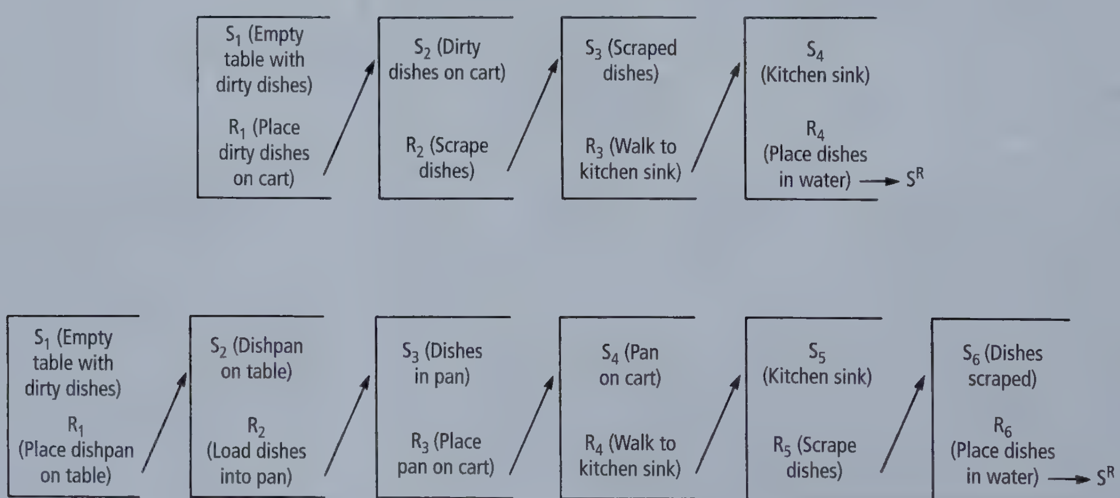


Figure 23.18 Illustration of an original behavior chain (upper panel) and its revision (lower panel) to break an inappropriate response sequence.

behaviors. Specifically, in the presence of dirty dishes on the empty table, the individual scraped excess food scraps onto the table itself, instead of placing the dirty items on the food cart. In other words, the initial S^D for the table-busing chain, the empty table with the dirty plates, was setting the occasion for a response (dish scraping) that should have occurred later in the chain.

To troubleshoot this inappropriate chain and determine a course of action, the behavior analyst should consider several possible sources of difficulty, including (a) re-examining the S^Ds and responses S_1-R_1-S_2-R_2; (b) determining whether similar S^Ds cue different responses; (c) analyzing the job setting to identify relevant and irrelevant S^Ds; (d) determining whether S^Ds in the job setting differ from training S^Ds; and (e) identifying the presence of novel stimuli in the environment.

Re-examine the S^D and Response Sequence

Kuhn et al. (2006) indicated that appropriate chains can be broken inadvertently and the entire chain, or key links within the chain, can be compromised when parents, teachers, and/or caregivers unwittingly reinforce an inappropriate sequence. They stated,

> . . . if a series of appropriate behaviors are inadvertently linked in a chain and the contingencies change for just some members of the chain (e.g., if a new caregiver begins to reinforce the terminal response whether or not it followed earlier responses in the chain), earlier appropriate behaviors could be eliminated (p. 265).

So, by re-examining the S^D and response sequence, a more appropriate arrangement might be contrived.

Determine Whether Similar S^Ds Cue Different Responses

Figure 23.18 (upper panel) shows two similar S^Ds—dirty dishes on the empty table (S_1) and dirty dishes on the cart (S_2)—which might have contributed to the scraping of food onto the table. In other words, R_2 (scrape dishes) might have come under the control of S_1 (dirty dishes on table). Figure 23.18 (lower panel) shows how the behavior analyst corrected the sequence by rearranging the S^Ds and their associated responses. Scraping the dishes is now the fifth response in the chain (R_5), and it occurs at the kitchen sink, an area away from the restaurant tables and diners. Thus, any potential confusion would be reduced or eliminated.

Analyze the Natural Setting to Identify Relevant and Irrelevant S^Ds

The training program should be designed so that the learner is trained to discriminate the relevant (i.e., critical) components of a stimulus from the irrelevant variations. Figure 23.18 (lower panel) shows at least two relevant characteristics of S_1: an empty table and the presence of dirty dishes on that table. An irrelevant stimulus might be the location of the table in the restaurant, the number of place settings on the table, or the arrangement of the place settings. A relevant characteristic of S_5, the presence of the kitchen sink, would be water faucets, the sink configuration, or dirty dishes. Finally, irrelevant stimuli might be the size of the kitchen sink and the type or style of faucets.

Determine Whether S^Ds in the Natural Setting Differ from Training S^Ds

It is possible that variations of the S^Ds cannot be taught during simulated training phases because natural conditions cannot be replicable or governmental regulations are prohibitive. For example, instrument pilots executing a complex chain of approach and landing maneuvers for a "circling approach" to an airport as part of a competency check, can only do that maneuver in a real airplane at an airport. Simulated approaches are not allowed. Hence, it is recommended that the final training session, at least for selected chains, be conducted in the natural setting where the behavior chain is expected to be performed. This allows any differences that exist to be recognized by the trainer and subsequent discrimination training to be further refined on site.

Identify the Presence of Novel Stimuli in the Environment

The presence of a novel stimulus unexpected in the original training situation can also prompt the occurrence of an inappropriate chain. In the restaurant example, the presence of a crowd of customers may set the occasion for the chain to be performed out of sequence. Likewise, distracting stimuli (e.g., customers coming and going, tips left on the table) may set the occasion for an inappropriate chain. Also, a coworker might unwittingly give contradictory instructions to the trainee. In any of these situations, the novel stimuli should be identified and the learner taught to discriminate them along with other S^Ds in the environment.

FACTORS AFFECTING THE PERFORMANCE OF BEHAVIOR CHAINS

Several factors affect the performance of a behavior chain. The following sections outline these factors and provide recommendations for addressing them.

Completeness of the Task Analysis

The more complete and accurate the task analysis, the more progress the person can make through the sequence. If the elements making up the chain are not sequenced appropriately, or if the corresponding S^Ds are not identified for each response, learning the chain will be more difficult.

The behavior analyst should remember two key points when attempting to develop an accurate task analysis. First, planning must occur before training. The time devoted to constructing and validating the task analysis is well spent. Second, after the task analysis is constructed, training should begin with the expectation that adjustments in the task analysis, or the use of more intrusive prompts, may be needed at various steps in the task analysis. For example, McWilliams and colleagues (1990) noted that one of their students required extensive trials and more intrusive prompts with a few steps in the task analysis before improved performance was obtained.

Length or Complexity of the Chain

Longer or more complex behavior chains take more time to learn than shorter or less complex behavior chains. Likewise, if two or more chain clusters are linked, practitioners can anticipate that training will be longer.

Schedule of Reinforcement

When a reinforcer is presented subsequent to the performance of a behavior in a chain, it affects each of the responses making up the chain. However, the effect on each response is not identical. For example, in backward chaining, responses performed at the end of the chain are strengthened faster than responses made earlier in the chain because they are reinforced more frequently. The behavior analyst is advised to remember two points: (a) A chain can be maintained if an appropriate schedule of reinforcement is used (see Chapter 13), and (b) the number of responses in a chain may need to be considered when defining the schedule of reinforcement.

Stimulus Variation

Bellamy, Horner, and Inman (1979) provided an excellent pictorial representation of how stimulus variation affects a learner's performance of a chain. The upper photograph in Figure 23.19 shows a cam switch bearing before and after its placement on a cam switch axle. The lower photograph shows the four different types of bearings that can be used in the assembly process. Any bearing sharing two critical features—a 1.12-cm circular hole in the center and one or more hex nut slots on its face—should evoke placing it on a cam switch axle. Bearings without these features, regardless of the presence of irrelevant features (e.g., color, material composition, weight), should not occasion the response.

If possible, the behavior analyst should introduce all possible variations of the S^D that the learner will encounter. Regardless of the behavior chain, presentation of stimulus variations increases the probability that the correct response will occur in their presence. This includes, for example, in an assembly task, various canisters and shafts; with dressing skills, different fasteners, zippers, and buttons; and with tooth brushing, assorted toothpaste tubes and pumps and brushes.

Response Variation

Often when stimulus variations occur, response variation also must occur to produce the same effect. Again, Bellamy and colleagues (1979) provided an illustration with the cam shaft assembly. In the upper-left photograph of Figure 23.20, the bearing has been placed on the cam shaft, and the retaining clip is being positioned for placement with a pair of nose pliers.

Figure 23.19 Upper panel: A cam switch bearing before and after placement on a cam switch axle. Lower panel: Four different types of bearings used in cam switch assembly.

From *Vocational Habilitation of Severely Retarded Adults*, pp. 40 & 42, by G. T. Bellamy, R. H. Horner, and D. P. Inman, 1979, Austin, TX: PRO-ED. Copyright 1979 by Pro-Ed. Reprinted by permission.

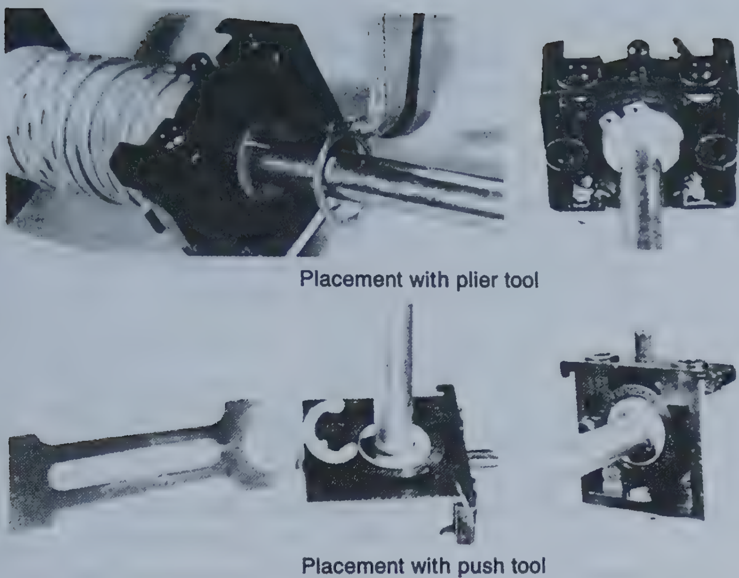


Figure 23.20 Two ways that retaining rings are applied to affix bearings to cam axles.

From *Vocational Habilitation of Severely Retarded Adults*, p. 44, by G. T. Bellamy, R. H. Horner, and D. P. Inman, 1979, Austin, TX: PRO-ED. Copyright 1979 by Pro-Ed. Reprinted by permission.

The upper-right photograph shows the clip in position. The lower-left photograph shows a different bearing configuration (i.e., the S^D is different) requiring a different response. Instead of the retaining clip being lifted over the bearing cap with pliers, it must be pushed over the cap with a wrench-type tool. The response of lifting or pushing has changed, as has the response of selecting the appropriate tool. Thus, the behavior analyst should be aware that when stimulus variation is introduced, training, or retraining of responses, within the chain may be required.

SUMMARY

Behavior Chain Defined

1. A behavior chain is a linked sequence of responses leading to a terminal outcome. Each response, and its associated stimulus condition, produces a change that functions as conditioned reinforcement for that previous response and as a discriminative stimulus for the next response in the chain. Reinforcement for the last response in a chain maintains the reinforcing effectiveness of the stimulus changes produced by all previous responses in the chain.

2. Each stimulus that links two sequential responses in a chain serves dual functions: It is a conditioned reinforcer for the response that produced it and an S^D for the next response in the chain.

3. A behavior chain has three important characteristics: (a) It involves the performance of a specific series of discrete responses; (b) the performance of each behavior in the sequence produces a stimulus change in the environment in such a way that it yields conditioned reinforcement for the preceding response and serves as an S^D for the next response; and (c) the responses within the chain must be performed in a specific sequence, and in close temporal succession.

4. In a behavior chain with a limited hold, a sequence of behaviors must be performed correctly and within a specific time for reinforcement to be delivered. Proficient responding is a distinguishing feature of chains with limited holds.

5. If a learner takes too much time to complete one or more component responses, the behavior analysis can implement contingencies focusing on the proficiency of performance. Setting a time criterion for the onset and completion of each response within the chain is one way to do this; another is to set a time criterion for completion of the total chain.

Rationale for Chaining

6. Whereas a behavior chain denotes a particular sequence of stimuli and responses ending in reinforcement, chaining refers to various methods for linking specific sequences of stimuli and responses to form new performances.

7. Chaining refers to the way these specific sequences of stimuli and responses are linked to form new performances. In forward chaining, behaviors are linked beginning with the first behavior in the sequence. In backward chaining, the behaviors are linked beginning with the last behavior in the sequence. Forward and backward chaining have variations available.

8. A behavior analyst should be skilled in building behavior chains for three reasons: (a) Chains can be used to improve independent-living skills; (b) chains can provide the means by which other behaviors can be combined into more complex sequences; and (c) chains can be combined with other procedures to build behavioral repertoires in generalized settings.

Establishing Behavior Chains with Task Analysis

9. Task analysis involves breaking a complex skill into smaller, teachable units, the product of which is a series of sequentially ordered steps or tasks.

10. The purpose of constructing and validating a task analysis is to determine the sequence of stepwise critical behaviors that constitute the complete task and that would lead to it being individualized for the learner and performed efficiently. Task analyses can be constructed by observing a competent performer, by executing the task yourself, by asking an expert, and by trial and error.

11. The purpose of assessing mastery level is to determine which components of the task analysis can already be performed independently. Assessment can be conducted with the single-opportunity or multiple-opportunity method.

Behavior Chaining Methods

12. In forward chaining, the behaviors identified in the task analysis are taught in their natural sequence. Specifically, reinforcement is delivered when the predetermined criterion for the first behavior in the sequence is achieved. Thereafter, reinforcement is delivered for criterion completion of Steps 1 and 2. With each successive step, reinforcement is delivered contingent on the correct performance of all steps trained thus far.

13. Total-task chaining is a variation of forward chaining in which the learner receives training on each step in the task analysis during every session. Trainer assistance is provided using response prompts with any step the individual is not able to perform. The chain is trained until the learner performs all the behaviors in the sequence to criterion.

14. In backward chaining, all the steps identified in the task analysis are completed by the trainer, except the last one. When the final step in the sequence is performed to criterion, reinforcement is delivered. Next, reinforcement is delivered when the next-to-last step and the last step are performed. Subsequently, the individual must perform the last three steps before reinforcement is delivered, and so on. The primary advantage of backward chaining is that the learner contacts the chain's terminal reinforcer on every instructional trial.

15. Backward chaining with leap aheads follows essentially the same procedures as in backward chaining, except that not every step in the task analysis is trained. The leap-aheads modification provides for probing or assessing untrained behaviors in the sequence. Its purpose is to speed up training of the behavior chain.

Choosing a Chaining Method

16. The decision to use forward chaining, total-task chaining, backward chaining, or backward chaining with leap aheads should be based on the results of a task analysis assessment; empirically sound data-based studies; and a functional evaluation, taking into consideration the cognitive, physical, motoric, and attention abilities of the individual.

Disrupting and Breaking Behavior Chains

17. Behavior chains can be interrupted or broken in at least six ways. These include extinction, satiation, unchaining, interrupting the chain, substituting an S^D, or extending the chain and building in time delays.

Troubleshooting Chains

18. To troubleshoot an inappropriate chain, the behavior analyst should consider (a) re-examining the S^Ds and responses S_1-R_1-S_2-R_2, (b) determining whether similar S^Ds cue different responses, (c) analyzing the job setting to identify relevant and irrelevant S^Ds, (d) determining whether S^Ds in the job setting differ from training S^Ds, and (e) identifying the presence of novel stimuli in the environment.

Factors Affecting the Performance of Behavior Chains

19. Factors affecting the performance of a behavior chain include (a) completeness of the task analysis, (b) length or complexity of the chain, (c) schedule of reinforcement, (d) stimulus variation, and (e) response variation.

KEY TERMS

backward chaining
backward chaining with leaps ahead
behavior chain
behavior chain interruption strategy
behavior chain with a limited hold
chaining
forward chaining
task analysis
total-task chaining

MULTIPLE-CHOICE QUESTIONS

1. Each of the following is an example of a behavior except:
   a. Mowing the lawn
   b. Brushing your teeth
   c. Blinking your eye
   d. Writing your name
   Hint: (See "Behavior Chain Defined")
2. A behavior chain with a limited hold means:
   a. The number of steps in the chain is limited.
   b. The learner must "hold" his/her responses for a specified period of time.
   c. The entire chain or steps within the chain must be completed in a given period of time.
   d. Accuracy of the performance of steps is not important.
   Hint: (See "Behavior Chain Defined")
3. Behavior chains and chaining procedures are important because:
   a. There are many behaviors that are considered behavior chains, and these behaviors lead to more independence in life.
   b. Chaining can be combined with many other teaching procedures so that behaviors can be taught in a variety of contexts.
   c. Chaining is an efficient training procedure-it can be used to combine already existing repertoires of behavior into new behaviors.
   d. All of these
   Hint: (See "Rationale for Chaining")
4. How are forward and backward chaining different?
   a. Forward chaining means the learner learns the steps to the skill in the correct order, and backward chaining means the learner learns the steps in the reverse order.
   b. Forward chaining means the learner learns the first step first in the sequence, and backward chaining means the learner learns the last behavior in the sequence first (but the steps are still learned in the correct sequence).
   c. Forward chaining is a much quicker way to learn a task than backward chaining.
   d. None of these answer choices accurately differentiates between forward and backward chaining.
   Hint: (See "Rationale for Chaining")
5. If you want to create a task analysis for a behavior you have no idea how to do yourself, how could you construct it?
   a. You could find people who are good at the behavior and ask them how to do it.
   b. You could watch people who know how to do the task and observe what they do.
   c. You could go to the literature to see if there are any published task analyses of the behavior.
   d. All of these.
   Hint: (See "Establishing Behavior Chains with Task Analysis")
6. What is a risk of choosing the single-opportunity method of assessment for assessing mastery of a task analysis?
   a. You may underestimate the learner's abilities.
   b. You may inadvertently be teaching the individual the behavior during assessment.
   c. You may frustrate the learner because of the time-consuming nature of this form of assessment.
   d. All of these.
   Hint: (See "Establishing Behavior Chains with Task Analysis")
7. Breaking down longer chains into smaller chains and/or skill clusters, teaching each skill cluster, and then chaining each skill cluster is a variation of which form of behavior chaining?
   a. Forward chaining
   b. Backward chaining
   c. Total task chaining
   d. Backward chaining with leap aheads
   Hint: (See "Behavior Chaining Methods")
8. Peer models have been used to teach children with disabilities to perform complex tasks using which form of chaining?
   a. Forward chaining
   b. Backward chaining
   c. Total task chaining
   d. Backward chaining with leap aheads
   Hint: (See "Behavior Chaining Methods")
9. The behavior chain interruption strategy (BCIS) can be used to:
   a. Teach a complex new behavior chain
   b. Increase speech, picture communication, and microswitch activation
   c. Replace the initial SDs in a chain
   d. All of these
   Hint: (See "Disrupting and Breaking Behavior Chains")

10. Kara has been learning to board a public bus, put money in the slot, and sit in her seat, waiting for her stop. Then, she pushes the button to indicate when she wants to get off. She has practiced this skill in her small hometown, where there are rarely more than five people on the bus. She traveled to Chicago with her parents to visit her aunt. She rode the bus with her parents in Chicago and was unable to do many steps in the chain. For example, she wasn't able to put her money in the slot, and she had difficulty choosing a seat. Her parents noticed that she was very distracted by all the people on the bus, as compared to the few people who are on the bus at home. Which of the following is the most likely reason that Kara wasn't able to perform steps in the chain she had learned so well at home?

a. The S^Ds on the bus were too similar to S^Ds on the bus at home, and so they triggered different responses.

b. The steps of the task were different on this bus than on the one at home.

c. There were novel stimuli present in Chicago that weren't present at home and that interfered with Kara's chain.

d. All of these are reasonable explanations for Kara's trouble.

Hint: (See "Disrupting and Breaking Behavior Chains")

ESSAY-TYPE QUESTIONS

1. Define and give an example of a behavior chain.

   Hint: (See "Behavior Chain Defined")

2. Explain why it is important to teach behavior chains to people.

   Hint: (See "Rationale for Chaining")

3. Assume you want to conduct a task analysis of playing a video game. What are the three different ways you could approach developing that task analysis?

   Hint: (See "Establishing Behavior Chains with Task Analysis")

4. Explain how forward, backward, and total-task chaining are conducted.

   Hint: (See "Behavior Chaining Methods")

5. If Marcus tends to have pizza binges when he is stressed and wants to stop this behavior chain, how could he go about it?

   Hint: (See "Disrupting and Breaking Behavior Chains")

NOTES

1. See Chapter 17 for prompting procedures.
2. To see Barnabus in action, refer to *Barnabus: A study in complex behavior chaining* [Film]. (1979). Boston: Northeastern University, 1979.

PART 9

Decreasing Behavior with Nonpunishment Procedures

Part 9 describes nonpunishment interventions to decrease or eliminate problem behaviors. Chapter 24 details extinction procedures for problem behaviors linked to behavior maintained by positive reinforcement, negative reinforcement, and automatic reinforcement. This chapter includes sections on extinction effects, variables affecting resistance to extinction, guidelines for using extinction effectively, and when not to use extinction. Chapter 25 describes the four most frequently researched variations of differential reinforcement for decreasing problem behavior: differential reinforcement of (a) incompatible behavior (DRI), (b) alternative behavior (DRA), (c) other behavior (DRO), and (d) low rates of responding (DRL). Examples of applications and guidelines for the effective use of each of these differential reinforcement procedures are provided. Chapter 26 defines, gives examples, and offers guidelines for treating problem behavior by changing the environment before the behavior occurs. Antecedent interventions directed toward problem behaviors are described with respect to function, stimulus control, and contingency independence.

CHAPTER 24

Extinction

LEARNING OBJECTIVES

- State the definition of extinction.
- Compare and contrast the functional and procedural forms of extinction.
- Identify three misuses of the term extinction.
- Compare and contrast the use of extinction with behaviors maintained by positive, negative, and automatic reinforcement.
- Plot and label an extinction curve on a graph.
- Discuss the phenomenon of spontaneous recovery.
- Explain the effect of continuous and intermittent reinforcement on behavior's resistance to extinction.
- Discuss how establishing operations affect resistance to extinction.
- List several behavioral examples that demonstrate how response effort influences a behavior's resistance to extinction.
- List and discuss the ten guidelines for the use of extinction.

> Pavlov had studied a process that was in a sense the reverse of conditioning and took place more slowly. He called it "extinction." In my early notes I sometimes called it "adaptation," because it resembled the slow disappearance of the unconditioned response to a click. My first extinction curve turned up by accident. A rat was pressing the lever in an experiment on satiation when the pellet dispenser jammed. I was not there at the time, and when I returned I found a beautiful curve. The rat had gone on pressing although no pellets were received, at first more rapidly than usual since no time was lost in eating, but then more and more slowly as time wore on. Some oscillation between high and low rates made the cumulative record wavelike.
>
> The change was more orderly than the extinction of a salivary reflex in Pavlov's setting, and I was terribly excited. It was a Friday afternoon and there was no one in the laboratory whom I could tell. All that weekend I crossed streets with particular care and avoided all unnecessary risks to protect my discovery from loss through my death.
>
> In my new laboratory I collected more extinction curves, all with roughly the same wavelike form—the rats pressing rapidly at first but then more and more slowly until they stopped.
>
> —B. F. Skinner (1979, p. 95)

Withholding reinforcement for a previously reinforced behavior decreases the rate of that behavior: a tactic, process, and principle known as *extinction*. As a behavior change tactic, extinction provides zero probability of reinforcement. As a behavioral process, extinction refers to the diminishing rate of a previously reinforced behavior when that behavior no longer produces reinforcement. As a principle, extinction refers to the functional relation between withholding reinforcement for a previously reinforced behavior and the resultant diminishing rate of response.

Practitioners have used extinction effectively in a wide variety of settings with problem behaviors ranging from mild disruptions to aggressive and self-injurious behavior. The effectiveness of extinction depends primarily on identifying and consistently withholding the reinforcer maintaining the target behavior. Although executing an extinction procedure would appear simple, implementing it effectively in applied settings can be difficult. This chapter describes three function-based extinction tactics, addresses common misuses of the term extinction, identifies secondary effects of extinction, and explains variables affecting resistance to extinction. The chapter concludes with guidelines for using extinction effectively and a discussion of when not to use extinction as a sole behavior reduction tactic.

EXTINCTION DEFINED

Extinction as a behavior change tactic occurs when reinforcement of a previously reinforced behavior is discontinued; as a result, the occurrence of that behavior decreases in the future.[1] Keller and Schoenfeld (1950) defined extinction this way: "Conditioned operants are extinguished by severing the relation between the act and the effect. . . . The principle of Type R [operant] extinction may be put as follows: The strength of a conditioned operant may be diminished by withholding its

reinforcement" (pp. 70–71). Similarly, Skinner (1953) wrote, "When reinforcement is no longer forthcoming, a response becomes less and less frequent in what is called 'operant extinction.'" (p. 69).

Note that an extinction procedure does not prevent the target behavior from occurring. Rather, extinction terminates the response–reinforcer relation by changing the environment so that the behavior no longer produces reinforcement.

When the third-grade teacher instructed the class to begin a written exercise, Moira frequently shouted, jumped out of her seat, and knocked her materials to the floor. Moira's teacher would ask her to calm down, tell her to return to her seat, explain patiently why her outbursts distracted her classmates, describe how she should behave, and offer to let Moira do another task. Moira would smile and agree to behave appropriately, but her annoying disruptions persisted. Suspecting that Moira disrupted the class to get his attention, the teacher decided he would offer Moira an alternative activity, but he would no longer ask her to calm down, tell her to return to her seat, explain matters, or describe better behavior. After 4 days, during which her disruptions produced no teacher attention, or attention from her classmates, Moira's disruptions decreased to zero.

Procedural and Functional Forms of Extinction

Vollmer and Athens (2011) noted, "There appears to have been a strong presumption in the early applied behavior analysis literature that problem behavior is always maintained by attention" (p. 319). As a result, many behavior analysts and practitioners relied on a procedural form of extinction (i.e., just ignore it, and it will go away) as treatment for problem behavior. Procedural extinction was used in the previous classroom example. The teacher suspected that Moira disrupted the class to get his attention and began ignoring her disruptions. In this scenario, Moira's problem behavior was maintained by positive reinforcement in the form of contingent teacher attention (e.g., explaining why the outbursts distracted classmates, describing how she should behave), and when the teacher consistently withheld that attention, the disruptions decreased. However, simply defaulting to "ignore it" as a method for applying extinction may not withhold the reinforcer maintaining the problem behavior.

Current best practice for treating problem behavior begins with an assessment to determine the behavior's function for the client (see Chapter 27). If extinction is indicated as a treatment option, its form is then matched to the function of the problem behavior. Had a functional assessment shown that Moira's disruptive behavior was maintained by socially mediated attention, her teacher's withholding attention following disruptions (i.e., not asking her to calm down and return to her seat, not explaining why her disruptions distracted the class, and not describing how she should behave) would have constituted a function-based application of extinction.

If a functional assessment revealed that Moira's outbursts were maintained by escape from instructional demands (i.e., negative reinforcement), teacher attention would not have constituted an extinction condition and most likely would have been ineffective. Extinction for negatively reinforced behavior requires the discontinuation of escape or avoidance from the stimulus conditions that evoke the behavior. In this case, Moira's disruptions would no longer enable her to avoid the workbook exercise in favor of a preferred activity.

The clarification between procedural and functional variations of extinction has led to more effective treatments (Lerman & Iwata, 1996a). When the form of extinction matches the function of the problem behavior, the intervention is usually effective (Vollmer & Athens, 2011). Detailed procedures for implementing function-based extinction are described later in this chapter.

Extinction: Misuses of a Technical Term

With the possible exception of *negative reinforcement, extinction* is the most misunderstood and misused technical term in applied behavior analysis. As a technical matter, *extinction* should be used only to identify (a) withholding the reinforcer for a previously reinforced behavior (the procedure), (b) a decreasing response rate under an extinction procedure (the process), or (c) the functional relation between withholding reinforcement for a previously reinforced behavior and the resultant diminishing rate of response (the principle). Four common misuses of extinction terminology are described in this section.

Using Extinction to Refer to Any Decrease in Behavior

A common error is using *extinction* to refer to any decrease in responding. Describing decreased responding that is due to any behavior change tactic or principle other than extinction (e.g., an abolishing operation, punishment, reinforcement for an alternative behavior) with a statement such as "the behavior is extinguishing" is conceptually and technically incorrect. Another common error is saying "the behavior extinguished" in reference to any reduction in behavior that reaches a zero count of occurrence.

Confusing Forgetting and Extinction

A misuse of the term *extinction* occurs when the speaker confuses extinction with forgetting. In forgetting, a behavior is weakened by the passage of time during which the person has no opportunity to emit the behavior. In extinction, behavior is weakened because it does not produce reinforcement.

Confusing Response Blocking and Sensory Extinction.

Applied behavior analysts have used objects such as goggles, gloves, helmets, and wrist weights to mask the automatic positive reinforcement received from sensory stimulation during sensory extinction. The applications of response blocking to reduce problem behaviors appear similar to sensory extinction. Response blocking, however, is not an extinction procedure. With all extinction procedures, including sensory consequences, the persons can emit the problem behavior, but that behavior will not produce reinforcement. By contrast, response blocking prevents the occurrence of the target behavior (Vollmer & Athens, 2011).

Confusing Noncontingent Reinforcement and Extinction

Noncontingent reinforcement (NCR) is another tactic for decreasing behavior that is sometimes confused with extinction.

With NCR, stimuli with known reinforcing properties are delivered on a fixed-time or variable-time schedule independent of responding (see Chapter 26). NCR does not entail withholding the reinforcer that maintains the problem behavior.

NCR is an important and effective intervention for diminishing problem behavior, but NCR operates on behavior in a different way than does extinction. Although both procedures produce a decrease in behavior, the behavioral effects result from different controlling variables. Extinction diminishes behavior by eliminating the reinforcing consequence; NCR diminishes behavior by creating an abolishing operation (see Chapter 16).

Misusing extinction terminology may be more harmful than a conceptual or technical language error. Such mistakes can lead to erroneous assumptions and faulty treatment decisions. Box 24.1: "Extinction Semantics" provides examples of correct, incorrect, and questionable statements inferring extinction with respect to treatments intended to decrease behavior, observed effects, and the principles responsible for the behavior change.

EXTINCTION PROCEDURES

Specifically, extinction procedures take three distinct forms that are linked to behavior maintained by positive reinforcement,[2] negative reinforcement, and automatic reinforcement.

Extinction of Behavior Maintained by Positive Reinforcement

Behaviors maintained by positive reinforcement are placed on extinction when those behaviors do not produce the reinforcer. In a classic extinction study, Williams (1959) described the effects of removing positive reinforcement on the tyrant-like behavior of a 21-month-old boy. The child had been seriously ill

BOX 24.1

Extinction Semantics

Statement	Meaning
We used *extinction* to treat Yoshiaki's inappropriate call-outs.	Refers to extinction as a *behavior change procedure*. The statement is correct if the procedure withheld the previous sources of reinforcement for Yoshiaki's call-outs. The statement is incorrect if the procedure withheld presumed sources of reinforcement for Yoshiaki's call-outs (e.g., "We ignored Yoshiaki's call-outs.").
Yoshiaki's call-outs are undergoing *extinction*.	Refers to extinction as a *behavioral process*. The statement is correct if describing a decreasing rate of response during a treatment condition in which previous sources of reinforcement for Yoshiaki's call-outs are withheld. The statement is incorrect if describing a decreasing rate of response during a treatment condition in which presumed sources of reinforcement for Yoshiaki's call-outs are withheld. If that is the case, the decreased rate of responding is due to other factors.
Extinction decreased Yoshiaki's call-outs.	Refers to extinction as a *principle of behavior*. The statement is correct only if an experimental analysis demonstrated a functional relation between the extinction procedure and the decreased occurrence of the behavior.
Audrey's second-language skills have *extinguished*. She has not spoken Italian since her last language class 3 years ago.	Attributing this lack of behavior to extinction is correct only if Audrey's attempts to speak Italian since her language class produced no reinforcement. If Audrey's Italian was weakened by the passage of time during which she had no opportunities to speak the language, then the statement confused extinction with forgetting.
Because a functional analysis suggests that Lefki's problem behavior is maintained by negative reinforcement, I recommend *escape extinction* as treatment. When you present a task to Lefki and she begins to engage in the problem behavior, physically guide her through the task, but do not let her escape from the task situation.	This statement correctly identifies escape extinction as the function-based form of extinction and describes a procedure for implementing it.
We have been *extinguishing* Jeremy's eye rubbing by blocking hand movements toward his eyes, and it's working.	This statement probably confuses response blocking with extinction. If Jeremy's eye rubbing was maintained by automatic reinforcement in the form of sensory stimulation, an extinction procedure would allow Jeremy to rub his eyes, but the responses would not produce the reinforcing stimulation.

for the first 18 months of his life but had completely recovered when the study began. The boy demanded special attention from his parents, especially at bedtime, and responded with temper tantrums (e.g., screaming, fussing, crying) when his parents did not provide the attention. A parent spent ½ to 2 hours each bedtime waiting in the bedroom until the boy went to sleep.

Although Williams did not speculate about how the tantrums developed, a likely explanation is not difficult to imagine. Because the child had been seriously ill for much of his first 18 months, crying had been his signal that he was experiencing discomfort or pain or needed help. In effect, his crying may have been reinforced by his parents' attention. Crying became a high-frequency behavior during the extended illness and continued after the child's health improved. The parents realized eventually that their child cried at bedtime to gain their attention and tried to ignore the crying. The crying increased in intensity and emotion when they did not stay in the room. As the days and weeks went by, the child's demands for attention at bedtime grew worse. The parents decided again not to stay with the child at bedtime; but the intensity of the crying increased, and some new tantrum behaviors occurred. The parents returned to the room and, by giving in to the tantrums, taught their child to behave like a tyrant.

After 3 months of the tyrant-like behavior, the parents decided they must do something about their child's tantrums. It appeared that parental attention maintained the tantrums; therefore, they planned to apply an extinction procedure. The parents, in a leisurely and relaxed manner, put the child to bed, left the bedroom, and closed the door. The parents recorded the duration of screaming and crying from the moment they closed the door.

The solid data path in Figure 24.1 shows the duration of tantrums for the 10 days following the first extinction intervention. The child had a tantrum for 45 minutes the first time he was put to bed without his parents staying in the room. The tantrums gradually decreased until the 10th session, when the child "no longer whimpered, fussed, or cried when the parents left the room. Rather, he smiled as they left. The parents reported that their child made happy sounds until he dropped off to sleep" (Williams, 1959, p. 269).

No tantrums occurred for approximately 1 week, but they resumed after his aunt put him to bed. When the child started to tantrum, the aunt returned to the bedroom and stayed with the child until he went to sleep. Tantrums then returned to the previous high level and had to be decreased a second time.

The dashed data path in Figure 24.1 during the second extinction intervention following the aunt's intervention. The data curve is similar to that of the first removal of parent attention. The duration of the tantrums was somewhat greater during the second removal, but reached zero by the ninth session. Williams reported that the child had no further bedtime tantrums during 2 years of follow-up.

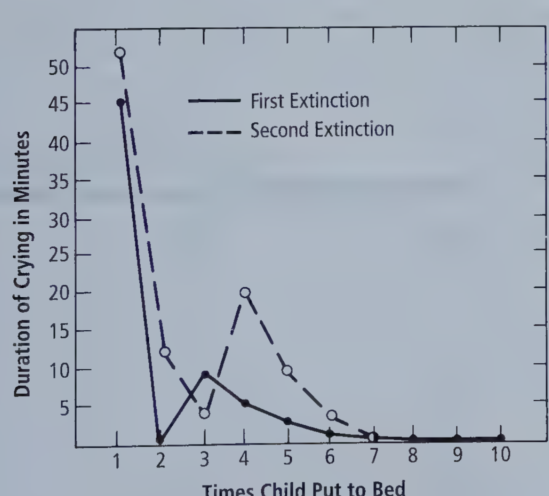


Figure 24.1 Two extinction series showing duration of crying as a function of being put to bed.

Based on "The Elimination of Tantrum Behavior by Extinction Procedures" by C. D. Williams, 1959, *Journal of Abnormal and Social Psychology*, *59*, p. 269. Copyright 1959 by the American Psychological Association.

Extinction of Behavior Maintained by Negative Reinforcement

Behaviors maintained by negative reinforcement are placed on extinction (also called **escape extinction**) when those behaviors do not produce a removal of the aversive stimulus, meaning that the person cannot escape from the aversive situation.

Anderson and Long (2002) provided treatment to reduce the problem behaviors of Drew, an 8-year-old boy with autism and moderate to severe intellectual disabilities. Drew's problem behaviors consisted of self-injurious behavior aggression, and disruptions. Results of a functional behavior assessment led Anderson and Long to hypothesize that escape from task situations maintained the problem behaviors. Based on this hypothesis, the speech therapist used escape extinction to diminish the problem behaviors that occurred while Drew worked on match-to-sample and receptive language tasks (these tasks evoked the highest rates of problem behavior). The speech therapist provided instructional prompts during the tasks. When Drew emitted problem behaviors following the instructional prompt, the speech therapist physically guided him to complete the task. Escape extinction produced a significant reduction in problem behaviors during the match-to-sample and receptive language tasks. Figure 24.2 shows the number of problem behaviors emitted per minute during baseline (i.e., escape) and escape extinction conditions.

A study by Dawson and colleagues (2003) provides another excellent example of escape extinction. They used escape extinction to help Mary, a 3-year-old whose total food refusal resulted in her admission to an outpatient program for the treatment of severe behavior problems. Her medical history included gastroesophageal reflux, delayed gastric emptying, and gastrostomy tube dependence, among other medical issues. Her food refusal included head turning when presented with a bite of food, putting her hand on the spoon or therapist's hand or arm, and using her hands or bib to cover her face.

The escape extinction procedure followed 12 sessions of food refusal. If Mary refused the bite of food, the therapist held the spoon to Mary's mouth until she took the bite. When Mary expelled the bite, the food was re-presented until swallowed. With this procedure, Mary's refusals did not produce an escape from the food. The therapist terminated the session after Mary

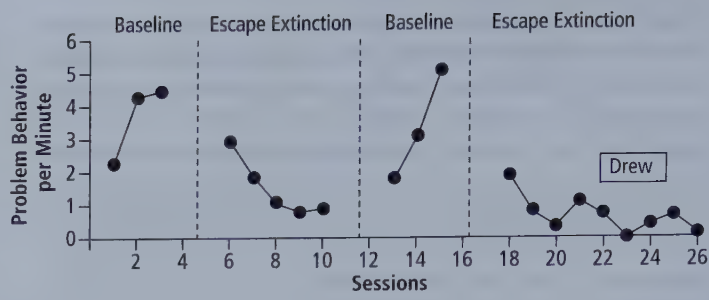


Figure 24.2 Rate of problem behaviors by Drew during baseline and escape extinction conditions.

Based on "Use of a Structured Descriptive Assessment Methodology to Identify Variables Affecting Problem Behavior" by C. M. Anderson & E. S. Long, 2002, *Journal of Applied Behavior Analysis*, *35*(2), p. 152. Copyright 2002 by Society for the Experimental Analysis of Behavior.

had swallowed 12 bites of food. Mary accepted no food during the 12 baseline sessions; but after 2 sessions of escape extinction, Mary's food acceptance increased to 100% compliance.

Extinction of Behavior Maintained by Automatic Reinforcement

Some behaviors produce natural sensory consequences that reinforce the behaviors that produce them. Rincover (1981) described a naturally occurring sensory reinforcer as a stimulus that "sounds good, looks good, tastes good, smells good, feels good to the touch, or the movement itself is good" (p. 1). Behaviors maintained by automatic positive reinforcement are placed on extinction (also called *sensory extinction*) by masking or removing the sensory consequence (Vollmer & Athens, 2011).

Automatic reinforcement can maintain self-injurious behaviors and persistent, nonpurposeful, repetitive self-stimulatory behaviors (e.g., flipping fingers, head rocking, toe walking, hair pulling, fondling body parts). Kennedy and Sousa (1995) reported that a 19-year-old man with profound disabilities had poked his eyes for 12 years, which resulted in visual impairment in both eyes. Eye poking was thought to serve a sensory stimulation function because it occurred most frequently when he was alone. When Kennedy and Sousa used goggles to mask contact with his eyes, eye-poking behavior decreased markedly.

Deaver, Miltenberger, and Stricker (2001) used extinction to decrease hair twirling. Hair twirling is a frequent precursor to hair pulling, a serious self-injury. Tina, at 2 years and 5 months of age, received treatment for hair twirling and pulling. A functional analysis documented that Tina did not twirl or pull hair to obtain attention, and that hair twirling most often occurred when Tina was alone at bedtime. The sensory extinction procedure consisted of Tina wearing thin cotton mittens on both hands at naptime during day care and during bedtime at home. Figure 24.3 shows that sensory extinction diminished hair twirling to near zero levels at home and at day care.

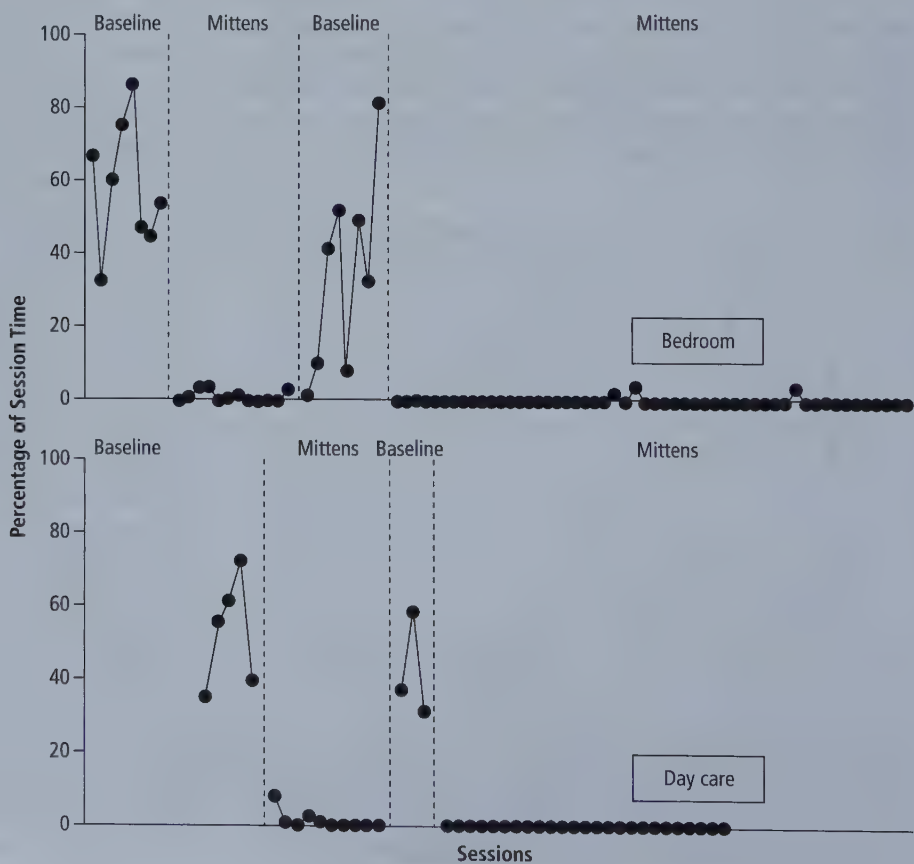


Figure 24.3 The percentage of session time of hair twirling during baseline and treatment conditions in the home and day-care settings.

Based on "Functional Analysis and Treatment of Hair Twirling in a Young Child" by C. M. Deaver, R. G. Miltenberger, and J. M. Stricker, 2001, *Journal of Applied Behavior Analysis*, *34*, p. 537. Copyright 2001 by Society for Experimental Analysis of Behavior.

Rincover, Cook, Peoples, and Packard (1979) and Rincover (1981) provided the following examples of applying extinction with automatic reinforcement:

1. A child persisted in flipping a light switch on and off. The visual sensory consequence was removed by disconnecting the switch.
2. A child persistently scratched his body until it bled. The tactile (touch) sensory consequence was removed by putting a thin rubber glove on his hand so that he could not feel his skin. Later, the glove was faded by gradually cutting off portions of it.
3. A child would throw up and then eat the vomit. The gustatory (taste) extinction procedure consisted of adding lima beans to the vomit. The child did not like lima beans; therefore, the vomit did not taste as good, and the positive sensory consequence was masked.
4. A child received kinesthetic stimulation (i.e., stimulation of muscles, tendons, and joints) by holding his arms out to his side and incessantly flapping his fingers, wrists, and arms. The extinction procedure consisted of taping a small vibratory mechanism on the back of his hand to mask the kinesthetic stimulation.
5. A child produced auditory stimulation by persistently twirling an object, such as a plate, on a table. Carpeting the surface of the table he used for spinning objects masked the auditory stimulation from his plate spinning.

SECONDARY EFFECTS OF EXTINCTION

When a previously reinforced behavior is emitted but not followed by the usual reinforcing consequences, the occurrence of that behavior will decrease to its pre-reinforcement level or stop entirely. However, numerous unreinforced responses can occur during an extinction procedure. The decrease in rate of response during extinction tends to be sporadic, with a gradual increase in pauses between responses.

In addition to extinction's primary effect of decreasing response rate, behaviors undergoing extinction often show other predictable characteristic effects, which can make the extinction procedure difficult for practitioners to apply: (a) extinction burst, (b) response variation, (c) initial increase in response magnitude, (d) spontaneous recovery, (e) resurgence, and (f) emotional outbursts and aggression. (Lattal, St. Peter Pipkin, & Escobar, 2013; Lerman & Iwata, 1995, 1996a; Murphy & Lupfer, 2014; Vollmer & Athens, 2011). These extinction effects have strong generality across species, response classes, and settings. Although applied behavior analysts almost always apply extinction as part of a treatment package that includes components that limit or prevent these extinction effects, practitioners must be aware of environmental conditions and competing contingencies that may exacerbate unwanted effects.

Extinction Burst

A common effect of the extinction procedure is an immediate increase in the rate of the response after removing the positive, negative, or automatic reinforcement. The behavioral literature uses the term **extinction burst** to identify this initial increase in rate of response. Figure 24.4 presents an illustration of an extinction burst. Operationally, Lerman, Iwata, and Wallace (1999) defined extinction burst as "an increase in responding during any of the first three treatment sessions above that observed during all of the last five baseline sessions or all of baseline" (p. 3). The extinction burst is well documented in basic research, but not well documented in applied research (Lerman & Iwata, 1995, 1996a). When reported, the bursts have occurred for only a few sessions without notable problems.

Goh and Iwata (1994) provided data showing an extinction burst. Steve was a 40-year-old man with profound developmental disabilities. He had been referred for evaluation for self-injury (head banging, head hitting). A functional analysis showed that self-injury was reinforced by escape from instructions. Goh and Iwata used extinction for the treatment of Steve's self-injury. The top panel in Figure 24.5 shows extinction bursts occurring with the onset of each of the two extinction phases.

Problem behaviors can worsen during extinction before they show improvement. For example, practitioners should anticipate an initial increase in disruption during extinction.

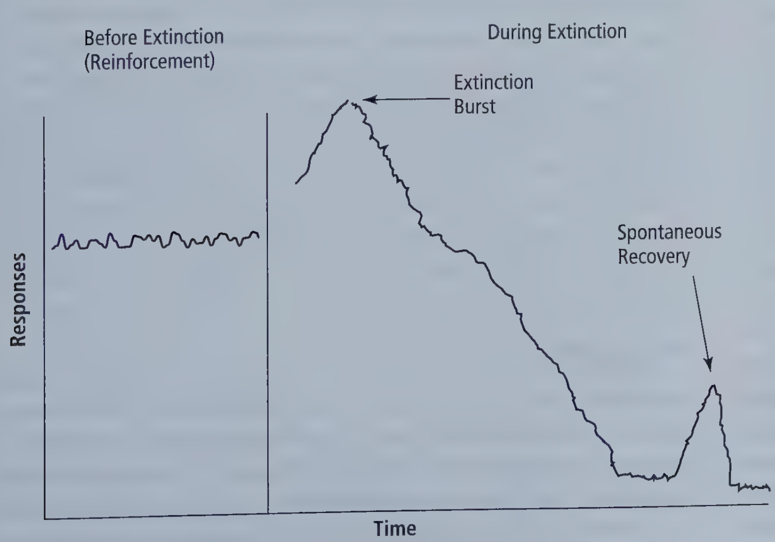


Figure 24.4 Hypothetical data illustrating an extinction burst and spontaneous recovery.

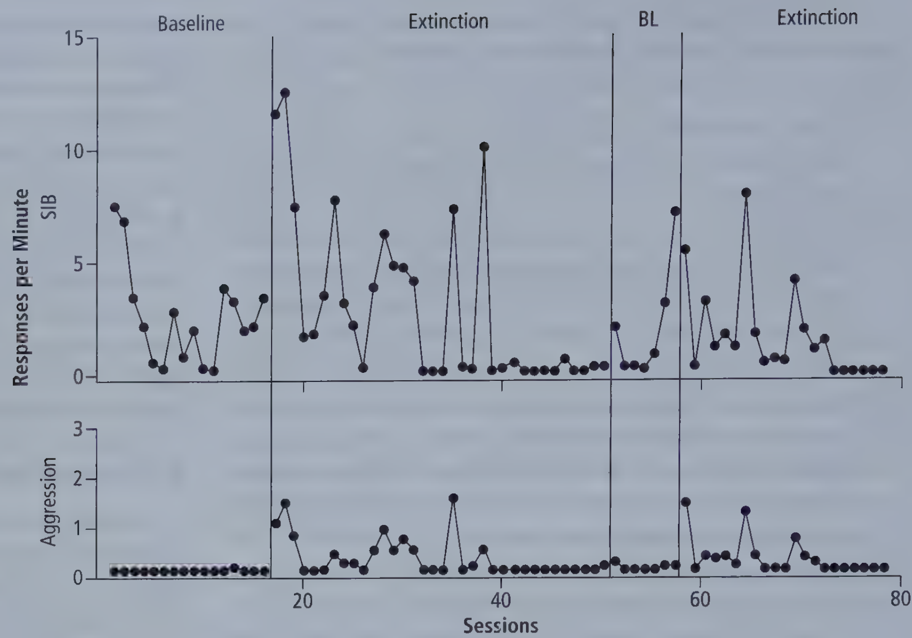


Figure 24.5 Number of responses per minute of self-injurious behavior (SIB) (top panel) and aggression (bottom panel) for Steve during baseline and extinction.

"Behavioral Persistence and Variability during Extinction of Self-Injury Maintained by Escape", by H.-L. Goh and B. A. Iwata, 1994, Reproduced with permission of John Wiley & Sons Inc.

Thereafter, problem behaviors should begin to decrease and should return to their pre-reinforcement level. Applied behavior analysts need to anticipate extinction bursts and be prepared to consistently withhold the reinforcing consequence. Extinction bursts usually suggest that the reinforcer (or reinforcers) maintaining the problem behavior was (or were) successfully identified, indicating that there is a good chance of an effective intervention.

Response Variation

Diverse and novel forms of behavior are sometimes observed during the extinction process (Kinloch, Foster, & McEwan, 2009; Peleg, Martin, & Holth, 2017). Commenting on this **extinction-induced variability**, Vollmer and Athens (2011) noted:

> From a layperson's perspective, it appears that the individual is "trying" to find new or at least other ways of obtaining reinforcement. For example, if a child asks politely for a reinforcer and the request is denied, he or she may begin to whine or attempt to "steal" the item. (p. 323)

Vollmer and Athens (2011) also raised an important consideration for applied behavior analysts. Although novel behaviors induced by extinction are often undesirable (e.g., whining, stealing, hitting), other novel behaviors can be desirable and appropriate. The behavior analyst can reinforce the appropriate novel behavior as an alternative behavior.

Initial Increase in Response Magnitude

An increase in response magnitude also may occur during the early stages of extinction. Parents who begin to ignore a child's bedtime tantrums might experience an increase in the loudness of screaming and the force of kicking before the tantrums begin to diminish. The extinction procedure is often difficult for teachers and parents to apply if there is an initial increase in response magnitude. For example, parents may be unwilling to ignore tantrum behavior for a sufficient amount of time because tantrums are aversive to parents and disruptive to other family members or siblings trying to go to sleep. Rolider and Van Houten (1984) presented a tactic for this practical problem. They suggested teaching parents to ignore gradually increasing durations of bedtime crying. They used baseline data to assess how long the parents could ignore bedtime crying comfortably before attending to their child. Then, the parents gradually increased their duration of ignoring. Every 2 days they waited an additional 5 minutes before attending to the child until a sufficient total duration of time was achieved.

Spontaneous Recovery

A behavior will typically continue a decreasing trend during extinction until it reaches a pre-reinforced level or ultimately ceases. However, a phenomenon commonly associated with the extinction process is the reappearance of the behavior after it has diminished to its pre-reinforcement level or stopped entirely. Basic researchers commonly report this extinction effect and call it **spontaneous recovery** (Rescorla, 2004). With spontaneous recovery, the behavior that diminished during the extinction process recurs even though the behavior does not produce reinforcement. Spontaneous recovery is short-lived and limited if the extinction procedure remains in effect (see Figure 24.4). Therapists and teachers need to know about spontaneous recovery, or they might conclude erroneously that the extinction procedure is no longer effective.

Resurgence

As Winterbauer, Lucke, and Bouton (2013) state: "An operant behavior that has undergone extinction can return ('resurge') when a second operant that has replaced it itself undergoes extinction" (p. 60). **Resurgence** refers to the reoccurrence of a previously reinforced behavior when reinforcement for an

alternative behavior is terminated or decreased and to the three-phase procedure that produces the effect: (1) A target behavior is reinforced, (2) the target behavior is placed on extinction and reinforcement provided for an alternative behavior, and (3) both responses are placed on extinction (Lattal & St. Peter Pipkin, 2009; Lieving & Lattal, 2003). During the final phase, the previously reinforced target behavior re-emerges, even though it remains under extinction condition.

The three-phase resurgence sequence might play out in an applied scenario as follows. (1) A student's rude remarks to other children on the playground produce reinforcement in the form of teacher attention. (2) The teachers supervising recess ignore the rude remarks (i.e., place the problem behavior on extinction) and attend to and praise all instances of appropriate interactions by the student (i.e., reinforce an alternative behavior). The intervention of extinction plus differential reinforcement of alternative behavior produces the desired effect: Occurrences of the student's inappropriate playground behavior decrease to zero and appropriate interactions increase. (3) Believing they have solved the problem, the teachers provide less and less attention to the alternative behavior. The resurgence of rude remarks by the student surprises the teachers.

Resurgence is not necessarily a negative phenomenon. Desirable resurgence occurs when a previously reinforced adaptive behavior recurs following an extinction period and during a period when other behaviors in the repertoire are unsuccessful (Epstein, 1991; Hoffman & Falcomata, 2014). Lattal and St. Peter Pipkin (2009) described a scenario in which the desirable resurgence of appropriate academic behavior might occur.

> Consider a situation in which a student taking an examination suddenly remembers information related to a topic. This could develop as follows. Assume that a teacher praises the use of a particular strategy to solve problems (similar to the reinforcement phase), but that this strategy is unsuccessful during subsequent lessons that involve structurally different problems (similar to the alternative reinforcement phase). With neither response now reinforced (extinction; for example, during the examination, when immediate praise would not be available for engaging in either response), the previously effective strategy may resurge. (pp. 255–256)

Although resurgence is often temporary, especially if those responses are not reinforced and some instances of the alternative behavior contact reinforcement (Wacker et al., 2010), it can also be a robust effect (Winterbauer et al., 2013). Basic and translational research has identified numerous factors that impact resurgence, including the target behavior's original reinforcement history prior to extinction (Winterbauer & Bouton, 2010), schedules of reinforcement for the alternative behavior (Winterbauer & Bouton, 2012), reinforcing the alternative behavior in a different context from that in which the problem behavior had previously been reinforced (Mace et al., 2010), reinforcing multiple alternative behaviors (Bloom & Lambert, 2015), response effort (Wilson, Glassford, & Koerkenmeier, 2016), the presence of discriminative stimuli during the resurgence test (King & Hayes, 2016), and lapses in treatment integrity (Fuhrman, Fisher, & Greer, 2016). To learn more about resurgence and its relevance for applied behavior analysts, see Lattal and St. Peter Pipkin (2009), St. Peter Pipkin (2015), and Kestner and Peterson (2017).

Emotional Outbursts and Aggression

Placing a behavior on extinction may occasion an increase in intensity or force of the response, and may evoke other emotional or aggressive behaviors (Vollmer & Athens, 2011). Common emotional outbursts and aggressive behaviors include escape, rage, prolonged whining, and crying. Other examples for the emotional outbursts and aggression may include behaviors such as escape behavior, rage, and crying.

VARIABLES AFFECTING RESISTANCE TO EXTINCTION

Behavior analysts refer to continued responding during an extinction procedure as **resistance to extinction**. Resistance to extinction is a relative concept. Behavior that continues to occur during extinction is said to have greater resistance to extinction than behavior that diminishes quickly. Three measures of resistance to extinction appear in the behavior analysis literature: (a) declining rate of response, (b) total count of responses emitted before responding ceases or attains some final low level, and (c) duration of time required for the behavior to reach a predetermined criterion (Lerman, Iwata, Shore, & Kahng, 1996; Reynolds, 1968).

Continuous and Intermittent Reinforcement

Continuous and intermittent schedules of reinforcement and their effects are described in Chapter 13. We can make three tentative statements about resistance to extinction as it relates to continuous and intermittent reinforcement. (a) Intermittent reinforcement (INT) may produce behavior with greater resistance to extinction than behavior previously reinforced by continuous reinforcement (CRF) (Keller & Schoenfeld, 1950). (b) Some intermittent schedules may produce more persistent responding during extinction than others (Ferster & Skinner, 1957). For example, variable schedules of reinforcement (variable ratio [VR], variable interval [VI]) may yield behavior more resistant to extinction than fixed schedules (fixed ratio [FR], fixed interval [FI]). (c) To a degree, the thinner the schedule of reinforcement, the greater the resistance to extinction.

Applied behavior analysts and practitioners should consider these statements as tentative guidelines, as researchers continue to discover schedule factors that influence extinction behavior. For example, recent density of reinforcement may predict continued responding during extinction (e.g., Lerman et al., 1996; Mace et al., 2011). MacDonald, Ahearn, Parry-Cruwys, and Bancroft (2013) found that four boys with autism persisted in problem behavior more so in 5-min extinction sessions that followed CRF than in extinction sessions immediately following INT reinforcement. Lerman et al. (1996) found that extinction bursts were more likely following CRF than INT.

Motivating Operations

All stimuli that function as reinforcers require a minimum level of an establishing operation (i.e., motivation must be present; see Chapter 16). The strength of the establishing operation (EO) above the minimum level will influence resistance to extinction. Basic research has shown that "resistance to extinction is greater when extinction is carried out under high motivation than under low" (Keller & Schoenfeld, 1950, p. 75).

Number, Magnitude, and Quality of Reinforcement

The number of times a behavior has produced reinforcement may influence resistance to extinction. A behavior with a long history of reinforcement may have more resistance to extinction than a behavior with a shorter history of reinforcement. If bedtime tantrums have produced reinforcement for 1 year, they might have more resistance to extinction than tantrums that have produced reinforcement for 1 week. Also, the magnitude and quality of a reinforcer will likely influence resistance to extinction. A reinforcer of greater magnitude and quality might produce more resistance to extinction than a reinforcer with less magnitude and quality.

Number of Previous Extinction Trials

Successive applications of conditioning and extinction may influence resistance to extinction. Sometimes problem behaviors diminish during extinction and then are accidentally strengthened with reinforcement. When this happens, applied behavior analysts can reapply the extinction procedure. Typically, behavior diminishes with fewer total responses during a reapplication of extinction. This effect is additive when the participant can discriminate the onset of extinction. With each successive application of extinction, decreases in behavior become increasingly rapid until only a single response occurs following the withdrawal of reinforcement.

Response Effort

Applied researchers have produced data on response effort and its effect on resistance to extinction (e.g., Lerman & Iwata, 1996a; Wilson et al., 2016). The effort required for a response apparently influences its resistance to extinction. A response requiring greater effort diminishes more quickly during extinction than a response requiring less effort.

USING EXTINCTION EFFECTIVELY

Numerous guidelines for using extinction effectively have been published, with most authors providing similar recommendations. Presented below are nine guidelines for the successful implementation of extinction.

Withhold All Reinforcers Maintaining the Problem Behavior

A first step in using extinction effectively is to identify and withhold all sources of reinforcement for the target behavior. The effectiveness of extinction depends on the correct identification of the consequences that maintain the problem behavior. Functional behavior assessments have improved greatly the application and effectiveness of using extinction in applied settings (Vollmer & Athens, 2011; Chapter 27).

Applied behavior analysts collect data on antecedent and consequent stimuli that are temporally related to the problem behavior and provide answers to questions such as the following:

1. Does the problem behavior occur more frequently when something happens in the environment to occasion the behavior (e.g., a demand or request)?
2. Is the frequency of the problem behavior unrelated to antecedent stimuli and social consequences?
3. Does the problem behavior occur more frequently when it produces attention from other persons?

If the answer to the first question is yes, the problem behavior may be maintained by negative reinforcement. If the answer to the second question is yes, the applied behavior analyst will need to consider withholding tactile, auditory, visual, gustatory, olfactory, and kinesthetic consequences alone or in combination. If the answer to the third question is yes, the behavior may be maintained by positive reinforcement in the form of social attention.

The consequence that maintains the problem behavior appears obvious in some applied settings. In the Williams (1959) study described earlier in this chapter, parental attention appeared to be the only source of reinforcement maintaining the tyrant-like behavior at bedtime. However, behaviors are frequently maintained by multiple sources of reinforcement. The class clown's behavior might be maintained by the teacher's reaction to the disruptive behavior, by attention from peers, or by a combination of both. When Johnny's parent brings him to preschool, he might cry to escape from preschool, to keep his parents with him, to occasion his teacher's concern and attention, or to achieve some combination of all three. When multiple sources of reinforcement maintain problem behavior, identifying and withholding one source of reinforcement may have minimal or no effect on behavior. If teacher and peer attention maintain inappropriate classroom behavior, then withholding only teacher attention may produce little change in the problem behavior. To effectively apply the extinction procedure, the teacher must withhold her attention and reinforce the other students in the room when they ignore their classmate's clownish behavior.

Withhold Reinforcement Consistently

When the reinforcing consequences have been identified, teachers must withhold them consistently. All behavior change procedures require consistent application, but consistency is essential for extinction; otherwise, the behavior may inadvertently be placed on an intermittent schedule of reinforcement, thereby making it more resistant to extinction. Teachers, parents, therapists, and applied behavior analysts often report

that consistency is the single most difficult aspect in using extinction. The error of not withholding reinforcement consistently negates the effectiveness of the extinction procedure, and this point cannot be overemphasized.

Combine Extinction with Other Procedures

Although extinction can be an effective singular intervention, it is rarely recommended as such (Vollmer & Athens, 2011; Vollmer et al., 1998). We also recommend that applied behavior analysts always consider combining extinction with other treatments, especially the reinforcement of alternative behaviors. Two reasons support this recommendation. First, the effectiveness of extinction may increase when it is combined with other procedures, especially positive reinforcement. By combining extinction with differential reinforcement of appropriate behaviors, for example, the applied behavior analyst alters the environment by reinforcing appropriate alternative behaviors and placing problem behaviors on extinction. During intervention, the extinction procedure should not reduce the overall number of positive consequences received by a participant. Second, differential reinforcement and antecedent procedures may reduce extinction bursts and aggression (Lerman et al., 1999).

Rehfeldt and Chambers (2003) combined extinction with differential reinforcement for the treatment of verbal perseveration maintained by positive reinforcement in the form of social attention. The participant was an adult with autism. Rehfeldt and Chambers presented social attention contingent on appropriate and non-perseveration speech, and applied attention extinction by not responding to the participant's inappropriate verbal behavior. This combined treatment produced an increase in appropriate verbal responses and a decrease in perseveration. These data suggest that contingencies of reinforcement might maintain the unusual speech of some individuals with autism.

The bedtime pass developed by Friman et al. (1999) as treatment for pediatric resistance to bedtime is another example of an intervention combining extinction with reinforcement for appropriate alternative behavior. Upon being put to bed, the child is given a small notecard exchangeable for one brief trip out of the bedroom for specific activity (e.g., use the bathroom, one more hug from Mom or Dad). The parents then return the child to bed without comment and ignore any further crying or calling out. In a systematic replication of the bedtime pass with four 3-year-olds, Freeman (2006) reported the intervention eliminated bedtime resistance without producing extinction bursts.

Use Instructions

Contingencies of reinforcement affect the future frequency of behavior automatically. It is not necessary for people to know, to describe, or even to perceive that contingencies are affecting their behavior. However, behaviors sometime diminish more quickly during extinction when practitioners describe the extinction procedure to clients. For example, teachers frequently provide small-group instruction while other students do independent seatwork. When seatwork students ask questions, they interrupt instruction. Many teachers correct this problem by placing question asking on extinction. They simply ignore student questions until after the end of small-group instruction. This tactic is often effective. The extinction procedure, however, tends to be more effective when teachers inform the class that they will ignore all student questions until after the end of small-group instruction.

Plan for Extinction-produced Aggression

Behaviors that occurred infrequently in the past may become more prominent during extinction. These extinction-produced behaviors are often emotional and aggressive (Lerman et al., 1999; Vollmer & Athens, 2011). Skinner (1953) interpreted the change in response topographies (e.g., side effects) as emotional behaviors, including the aggression that sometimes accompanies extinction.

Goh and Iwata (1994) provided a convincing demonstration showing how aggression-induced behaviors occurred when a target behavior was placed on extinction. Steve was a 40-year-old man with profound intellectual disabilities. He had been referred for evaluation for self-injury (head banging, head hitting). A functional analysis showed that self-injury was reinforced by escape from instructions. Goh and Iwata used extinction for the treatment of Steve's self-injury. Steve rarely slapped or kicked other persons (i.e., aggression) during the two baseline phases, but aggression increased with the onset of the two extinction phases (see Figure 24.5, bottom panel). Aggression essentially stopped by the end of each extinction phase when self-injury was stable and low, even though Steve's aggression remained untreated during the baseline and extinction conditions.

Applied behavior analysts need to manage aggressive behaviors when they occur as side effects of extinction. It is critical that extinction-produced aggression not produce reinforcement. Frequently, parents, teachers, and therapists react to the aggression with attention that may function as reinforcement of the extinction-produced aggression. For instance, a teacher decides to ignore Moira's questions while providing small-group instruction. When Moira interrupts the teacher with a question, the teacher does not respond. Moira then starts disrupting other seatwork students. To quiet Moira down, the teacher responds, "Oh, all right, Moira, what did you want to know?" In effect, the teacher's behavior may have reinforced Moira's interruptions during small-group instruction and the accompanying inappropriate disruptions of the other seatwork students.

At times, extinction-produced aggression takes the form of verbal abuse. Often, teachers and parents do not need to react to it. If extinction-produced aggression results in reinforcement, individuals simply use other inappropriate behavior, such as verbal abuse, to produce reinforcement. However, teachers and parents cannot and should not ignore some forms of aggression and self-injurious behavior. They need to know (a) that they can ignore some aggression, (b) when they need to intervene on the aggression, and (c) what they will do for the intervention.

Increase the Number of Extinction Trials

An extinction trial occurs each time the behavior does not produce reinforcement. Whenever possible, applied behavior analysts should increase the number of extinction trials for the problem behaviors. Increasing the number of trials improves efficiency by accelerating the extinction process (i.e., the cumulative series of instances in which the behavior–controlling variable–primary effect sequence occurs). Applied behavior analysts can increase extinction trials when increased occurrences of the problem behavior can be tolerated during intervention. For example, Keith's parents used an extinction procedure to reduce his tantrums. His parents noticed that Keith had tantrums most often when he did not get his way about staying up late, eating a snack, and going outside. For the purpose of the program, they decided to set up several additional situations each day in which Keith did not get his way. Keith was then more likely to emit the inappropriate behavior more often, thereby giving his parents more occasions to ignore it. As a result, his tantrums decreased in a shorter period than they would have if left at their usual occurrence.

Include Significant Others in Extinction

It is important that other persons in the environment not reinforce undesirable behavior. A teacher, for example, needs to share extinction plans with other people who might help in the classroom—parent volunteers, grandparents, music teachers, speech therapists, industrial arts specialists—to avoid their reinforcing inappropriate behaviors. For extinction to be maximally effective, meaning having few, if any, instances of extinction bursts, significant increases in response magnitude, and so forth, all persons in contact with the learner must withhold reinforcement for the problem behavior.

Maintain Extinction-decreased Behavior

Applied behavior analysts leave the extinction procedure in effect permanently for maintaining the extinction-diminished behavior. A permanent application of escape extinction and attention extinction is a preferred procedure. Applied behavior analysts can use a permanent application also with some sensory extinction procedures. For example, the carpeting on a table used for spinning objects could remain in place indefinitely. Some applications of sensory extinction appear inappropriate and inconvenient if kept permanently in effect. For example, requiring Tina to permanently wear thin cotton mittens on both hands at naptime during day care and during bedtime at home appears inappropriate and inconvenient (see Figure 24.3). In such a case, the applied behavior analyst can maintain treatment gains by gradually fading out the sensory extinction procedure—for example, cutting off 1 inch of the glove palm every 3 to 4 days until the glove palm is removed. The analyst may then remove the fingers and thumb one at a time to gradually remove the mittens.

Guard Against Unintentional Extinction

Desirable behaviors are often unintentionally placed on extinction. A beginning teacher confronted with one student who is on task and many other students who are not will probably direct most of his attention to the majority and will provide little or no attention to the student who is working. It is common practice to give the most attention to problems ("the squeaky wheel gets the grease") and to ignore situations that are going smoothly. However, behaviors must continue to be reinforced if they are to be maintained. Teachers must be vigilant to "catch them being good" and deliver reinforcement accordingly.

WHEN NOT TO USE EXTINCTION

Although most interventions designed to decrease behavior will include withholding reinforcement for the target behavior, we envision at least four circumstances when applied behavior analysts should not use extinction as a singular intervention.

The Behavior Is Harmful

Some behaviors are so self-injurious, harmful to others, or destructive to property that they must be controlled with the most rapid and humane procedure available. Extinction as a singular intervention is not recommended in such situations.

All Sources of Reinforcement Cannot Be Withheld

Practitioners and significant others may discontinue reinforcement for previously reinforced behavior at every instance. If, however, the target behavior contacts other sources of reinforcement, the extinction procedure is doomed at the outset. A classic case in point is the previous scenario of the class clown obtaining attention from classmates while his teacher dutifully ignores his antics.

A Rapid Reduction in Response Rate Is Required

Practitioners often do not have the luxury of time to change behavior gradually. A vocational student who talks back to supervisors on a community job site may get fired before an intervention relying on extinction eliminates the problem behavior.

Others Are Likely to Imitate the Problem Behavior

Some behaviors can be tolerated if only one person emits them for a while. Those same behaviors may be intolerable if several persons begin emitting them. An extinction-only intervention is ill advised if others are likely to imitate the behavior placed on extinction ("Hey, if Thad can toss papers, turn over desks, or scream expletives, so can we!").

SUMMARY

Extinction Defined

1. Withholding reinforcement for a previously reinforced behavior decreases the rate of that behavior: a procedure, process, and principle known as *extinction*.
   - As a behavior change procedure, extinction provides zero probability of reinforcement.
   - As a behavioral process, extinction refers to the diminishing rate of a previously reinforced behavior when that behavior no longer produces reinforcement.
   - As a principle, extinction refers to the functional relation between withholding reinforcement for a previously reinforced behavior and the resultant diminishing rate of response.
2. Functional behavior assessments (see Chapter 27) enable behavior analysts to distinguish clearly between the procedural variations of extinction (i.e., most commonly—"Just ignore it, and it will go away."), and the functional variations of extinction (i.e., identifying the specific reinforcer maintaining the behavior).
3. Behavior analysts should use *extinction* only to identify (a) withholding the reinforcer for a previously reinforced behavior (the procedure), (b) a decreasing response rate under an extinction procedure (the process), or (c) the functional relation between withholding reinforcement for a previously reinforced behavior and the resultant diminishing rate of response (the principle).
4. Extinction (a) does not refer to any decrease in behavior, (b) is not the same as forgetting, (c) is not response blocking, and (d) is not noncontingent reinforcement.

Extinction Procedures

5. Procedures for extinction take three distinct forms that are linked to behavior maintained by positive reinforcement, negative reinforcement, and automatic (positive and negative) reinforcement.
6. Behaviors maintained by positive reinforcement are placed on extinction when those behaviors do not produce the reinforcer.
7. Behaviors maintained by negative reinforcement are placed on extinction (escape extinction) when those behaviors do not produce a removal of the aversive stimulus, meaning that the individual cannot escape from the aversive situation.
8. Behaviors maintained by automatic positive reinforcement are placed on extinction (sensory extinction) by masking or removing the sensory consequence.
9. Behaviors maintained by automatic negative reinforcement are placed on extinction by not removing (escape extinction) the aversive sensory stimulation.

Secondary Effects of Extinction

10. Behaviors undergoing extinction are usually associated with effects in occurrence and topography of response.
11. Common effects of extinction include (a) extinction burst, (b) response variation, (c) initial increase in response magnitude, (d) spontaneous recovery, (e) resurgence, and (f) emotional outbursts and aggression.

Variables Affecting Resistance to Extinction

12. Behavior that continues to occur during extinction is said to have greater resistance to extinction than behavior that diminishes quickly. Resistance to extinction is a relative concept.
13. Intermittent schedules of reinforcement may produce behavior with greater resistance to extinction than the resistance produced by a continuous reinforcement schedule.
14. Variable schedules of reinforcement (e.g., VR, VI) may have more resistance to extinction than fixed schedules (e.g., FR, FI).
15. To a degree, the thinner the intermittent schedule of reinforcement is, the greater the resistance to extinction will be.
16. Resistance to extinction is likely to increase with the strength of the establishing operation (EO) for the reinforcer being withheld.
17. The number, magnitude, and quality of the reinforcer may affect resistance to extinction.
18. Successive applications of conditioning and extinction may influence the resistance to extinction.
19. The response effort will likely influence resistance to extinction.

Using Extinction Effectively

20. Ten guidelines for using extinction are as follows:
    - Withhold all reinforcers maintaining the problem behavior.
    - Withhold reinforcement consistently.
    - Combine extinction with other procedures.
    - Use instructions.

- Plan for extinction-produced aggression.
- Increase the number of extinction trials.
- Include significant others in the intervention.
- Guard against unintentional extinction.
- Maintain extinction-decreased behavior.

When Not to Use Extinction

21. Extinction should not be used as a singular intervention in the following circumstances: when the target behavior is harmful, when all sources of reinforcement cannot be withheld, when a rapid reduction in behavior is required, or when others are likely to imitate the problem behavior.

KEY TERMS

escape extinction	extinction burst	sensory extinction
extinction (operant)	resistance to extinction	spontaneous recovery

MULTIPLE-CHOICE QUESTIONS

1. A procedure in which reinforcement of a previously reinforced behavior is discontinued and as a result, occurrences of that behavior decrease in the future is:
   a. Punishment
   b. Extinction
   c. Reinforcement
   d. Establishing Operation
   Hint: (See "Definition of Extinction")

2. What is the probability of reinforcement when utilizing an extinction procedure?
   a. 0%
   b. 15%
   c. 50%
   d. Variable
   Hint: (See "Definition of Extinction")

3. What form of extinction is used when the reinforcer (or reinforcers) maintaining a behavior is withheld?
   a. Procedural
   b. Functional
   c. Standard
   d. Respondent
   Hint: (See "Procedural and Functional Forms of Extinction")

4. What are the common misuses of the term extinction?
   a. Confusing forgetting and extinction
   b. Confusing response blocking and sensory extinction
   c. Confusing noncontingent reinforcement and extinction
   d. All of these
   Hint: (See "Extinction: Misuses of a Technical Term")

5. Delivering a stimulus with known reinforcing properties to an individual on a fixed time or variable time schedule independently of responding is known as which procedure?
   a. Extinction
   b. Noncontingent reinforcement (NCR)
   c. Differential reinforcement
   d. Punishment
   Hint: (See "Extinction: Misuses of a Technical Term")

6. What is the main effect of an extinction procedure?
   a. Increase in behavior
   b. No change in behavior
   c. Behavior decreases and/or stops entirely
   d. The effects are unknown
   Hint: (See: "Secondary Effects of Extinction")

7. An immediate increase in frequency and/or amplitude of the response after the removal of the maintaining reinforcer is called
   a. Extinction burst
   b. Reinforcement
   c. Noncontingent reinforcement
   d. Punishment
   Hint: (See "Extinction Burst")

8. The reappearance of a behavior after it has diminished to its pre-reinforced level and/or stopped entirely is called:
   a. Extinction burst
   b. Spontaneous recovery
   c. Differential reinforcement
   d. Procedural extinction
   Hint: (See "Spontaneous Recovery")

9. Continued responding during the extinction process is known as:
   a. Resistance to extinction
   b. Extinction burst
   c. Reinforcement
   d. Spontaneous recovery

   Hint: (See "Variables Affecting Resistance to Extinction")

10. Resistance to extinction is greater when carried out under:
    a. Low motivation
    b. High motivation
    c. No motivation
    d. None of these

    Hint: (See "Establishing Operation")

11. What suggestions did the authors give for assuring the effectiveness of extinction?
    a. Increasing the number of extinction trials
    b. Withholding all reinforcers maintaining the problem behavior
    c. Withholding reinforcement consistently
    d. Using instruction
    e. All of these

    Hint: (See "Using Extinction Effectively")

12. When is it not appropriate to utilize an extinction procedure?
    a. When inappropriate behaviors are likely to be imitated by others
    b. When behaviors are harmful to self or others
    c. Both of these
    d. None of these

    Hint: (See "When Not to Use Extinction")

ESSAY-TYPE QUESTIONS

1. Identify and briefly discuss four common misuses of the term extinction.

   Hint: (See "Extinction: Misuses of a Technical Term")

2. Compare and contrast the use of extinction with behaviors maintained by positive, negative, and automatic reinforcement.

   Hint: (See "Extinction of Behavior Maintained by Positive Reinforcement")

3. Discuss the phenomenon of spontaneous recovery.

   Hint: (See "Spontaneous Recovery")

4. Explain the effect of continuous and intermittent reinforcement on behavior's resistance to extinction.

   Hint: (See "Continuous and Intermittent Reinforcement")

5. List several behavioral examples that demonstrate how response effort influences a behavior's resistance to extinction.

   Hint: (See "Response Effort")

NOTES

1. The term *extinction* is also used with respondent conditioned reflexes (see Chapter 2). Presenting a conditioned stimulus (CS) again and again without the unconditioned stimulus until the CS no longer elicits the conditioned response is called *respondent extinction*.

2. Extinction linked to behaviors maintained by positive reinforcement is sometimes called *attention extinction*, especially in the functional behavior assessment (FBA) literature, in which reinforcement by social attention is one of the conditions or hypotheses explored by FBA.

CHAPTER 25

Differential Reinforcement

LEARNING OBJECTIVES

- Define and give examples of differential reinforcement.
- Define, give examples of, and explain how to use differential reinforcement of incompatible/alternative behavior.
- Define, give examples of, and explain how to use differential reinforcement of other behavior.
- Define, give examples of, and explain how to use differential reinforcement of low rates of behavior.

Does Love or Differential Reinforcement Open Doors?

There is a story about Queen Victoria and Prince Albert, which, although charming, has not been verified as historical fact (Strachey, 1921). Whether true or not, the story illustrates the use of differential reinforcement.

As the story goes, shortly after their marriage, Prince Albert and Queen Victoria had a quarrel. Albert walked (or stomped, depending on the telling) out of the room and locked himself in his study. Victoria ran after him and pounded on the door. "Who's there?" called Albert. "The Queen of England, and she demands to be admitted." Albert did not respond. Victoria pounded again. One version of the story has it that Victoria shouted, "I am the Queen of England, Scotland, Wales and Ireland, the Empress of India and of the entire British Commonwealth, I am the Commander-in-Chief of all the British armed forces, and I order you hereby to open this door!!!" Albert still did not respond. More furious knocking was followed by a pause. Then there was a gentle tap. "Who's there?" Albert asked. "Your wife, Albert," the Queen replied. The prince immediately opened the door.

This story is told on various Internet sites to extol the virtues of love or to make the point that it is more important to resolve conflict through nonviolent than violent means. But for behavior analysts, it is a fairly simple example of Prince Albert using good old-fashioned differential reinforcement (of alternative behavior) on the Queen's behavior, although we must assume her calmer, more solicitous behavior actually occurred more often under similar circumstances. Moreover, the immediate increase in the magnitude of her behavior (i.e., the extinction burst) suggests that such behavior had been reinforced in the past, either by Albert, or, more likely, by others.

And what about the effect of her capitulation on his behavior? When they quarreled in the future, was he more likely to storm out of the room and hole himself up in his study? We don't know, but we might predict so.

—H. D. Schlinger (2007, p. 385)

Practitioners can choose from a wide range of procedures for decreasing or eliminating problem behaviors. Although interventions based primarily on extinction or punishment are often effective, unwanted side effects may occur. Maladaptive emotional behavior and higher-than-usual rates of responding are commonly observed when a behavior with a long and consistent history of reinforcement is placed on extinction. Escape, avoidance, aggression, and other forms of undesirable counter-control may be evoked by punishment. In addition to unwanted side effects, another limitation of extinction and punishment as primary methods for reducing problem behaviors is that neither approach strengthens or teaches adaptive behaviors with which individuals can attain reinforcers. Beyond the possibility of unwanted side effects and their lack of educative value, interventions that rely on extinction alone or punishment in any of its forms raise important ethical and legal concerns. To dispel these limitations and concerns, applied behavior analysts have developed procedures that use differential reinforcement to diminish or eliminate problem behaviors.

DIFFERENTIAL REINFORCEMENT DEFINED

Differential reinforcement entails reinforcing one response class and withholding reinforcement for another response class. When used as a reductive procedure for problem behavior, differential reinforcement consists of (a) providing reinforcement contingent on a behavior other than the problem behavior or the problem behavior occurring at a reduced rate and (b) withholding reinforcement for the problem behavior.[1] Differential reinforcement can be implemented with positive reinforcement or negative reinforcement.

As Cowdery, Iwata, and Pace (1990) noted,

> differential reinforcement does not involve extended interruptions of ongoing activities (e.g., time-out), contingent removal of positive reinforcers (e.g., response cost), or presentation of aversive stimuli (e.g., punishment). These characteristics make differential reinforcement the least intrusive of all behavior interventions and probably account for its widespread popularity. (p. 497)

Three well-researched tactics for using differential reinforcement for decreasing inappropriate behavior are differential reinforcement of alternative behavior (DRA), differential reinforcement of other behavior (DRO), and differential reinforcement of low rates (DRL). This chapter defines, presents examples, and suggests guidelines for the effective use of each of these differential reinforcement procedures for decreasing problem behavior.

DIFFERENTIAL REINFORCEMENT OF ALTERNATIVE BEHAVIOR (DRA)

A practitioner using **differential reinforcement of alternative behavior (DRA)** reinforces occurrences of a desirable alternative to the problem behavior and withholds reinforcement for the problem behavior. DRA has the dual effects of strengthening the alternative behavior and weakening problem behavior (see Petscher, Rey, & Bailey, 2009 for a review). DRA can be conceptualized as a concurrent schedule of reinforcement in which two response classes receive reinforcement at different rates: (a) the alternative behavior on a dense schedule of high-quality reinforcement and (b) the problem behavior on extinction or a very lean schedule of reinforcement. Because the concurrent schedule favors the alternative behavior, the client allocates more responding to the alternative behavior and less responding to the problem behavior, as described by the matching law (Borrero et al., 2010; see Chapter 13).

Dixon, Benedict, and Larson (2001) used DRA to treat the inappropriate verbal behavior of Fernando, a 25-year-old man with moderate developmental disabilities and a psychotic disorder. Inappropriate vocalizations included utterances not related to the current context, sexually inappropriate remarks, illogical placement of words within a sentence, and "psychotic" statements (e.g., "There's a purple moose on my head named Chuckles," p. 362). The researchers defined appropriate verbal behavior as any vocal utterance that did not meet the defining characteristics of inappropriate vocalization. A functional analysis revealed that the inappropriate verbal statements were maintained by social attention.

During the DRA condition, the researchers attended to Fernando's appropriate statements with 10 seconds of comments relevant to his statements and ignored his inappropriate verbal behavior. For example, if Fernando said something about an activity he liked to do, the experimenter told him it was interesting and hoped he could do it again soon. Following an initial phase of treatment with DRA, the researchers created a comparison "baseline" condition by reversing the contingencies for the two response classes: Inappropriate vocalizations resulted in attention and appropriate verbal behavior was ignored. Data from the B-A-B-A-B design shows that the DRA intervention reduced Fernando's inappropriate verbal behavior and increased appropriate utterances (Figure 25.1).

In most published studies evaluating DRA, the alternative behavior and the problem behavior are mutually exclusive response classes whose different topographies make it impossible for the participant to emit both behaviors at the same time. In the Dixon et al. (2001) study, Fernando could not utter appropriate and inappropriate verbalizations at the same time. Similarly, a learner cannot simultaneously be in and out of his seat (Friman, 1990), politely ask "Can I have that please?" and scream (Schlichenmeyer, Dube, & Vargas-Irwin, 2015), comply with a therapist's instruction to complete a task and destroy the materials needed for the task (Ringdahl et al., 2002), or put an inedible item in a trashcan and in her mouth (Slocum, Mehrkan, Peters, & Vollmer, 2017). When DRA is implemented by reinforcing a behavior that cannot occur simultaneously with the problem behavior, the procedure is sometimes called **differential reinforcement of incompatible behavior (DRI)**.

The effectiveness of DRA does not depend upon an incompatible alternative behavior. An alternative behavior that occupies the participant when he might ordinarily engage in the problem behavior may be sufficient. Ideally, however, the alternative behavior serves a greater purpose than time filler. Adaptive activities or skills are the most beneficial alternative behaviors. For example, a classroom teacher could assign two students who frequently argue with each other to work on a class project together and reinforce cooperative behaviors associated with project development. Working together on a class project is not incompatible with arguing; the two response classes could occur together. However, the two students are likely to argue less when engaged in cooperative behaviors.

When differential reinforcement uses escape from a task or demand situation as a reinforcer, the procedure is sometimes identified as *differential negative reinforcement of alternative behavior* (DNRA) (e.g., Marcus & Vollmer, 1995; Piazza, Moses, & Fisher, 1996; Vollmer & Iwata, 1992). Applying DNRA to reduce the occurrence of a problem behavior maintained by escape from a task or demand context consists of providing negative reinforcement of the alternative behavior in the form of brief periods of escape from the task and escape extinction for the problem behavior.

With DNRA, positive reinforcement for the alternative behavior is often provided as well. For example, Lalli, Casey, and Kates (1995) allowed participants to escape from a task for

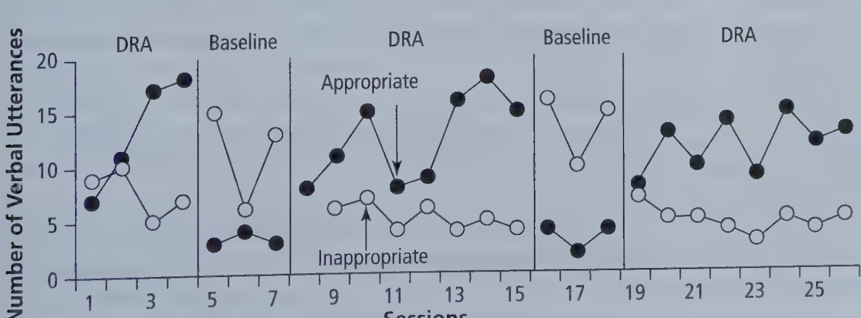


Figure 25.1 Number of appropriate and inappropriate verbal utterances by an adult male during DRA and baseline conditions.

"Functional Analysis and Treatment of Inappropriate Verbal Behavior" by M. R. Dixon, H. Benedict, & T. Larson (2001), *Journal of Applied Behavior Analysis*, *34*, p. 362. Reproduced with permission of John Wiley & Sons Inc.

30 seconds contingent on using an alternative response (e.g., giving the therapist a card with "BREAK" printed on it, or saying "no") as an alternative to their aberrant behavior. The researchers also provided praise for using the alternative responses. As training progressed, the participants' escape from a task was contingent on using the trained verbal responses and completing a gradually increasing number of required steps of the task. The DNRA intervention increased the use of the trained verbal response and decreased the problem behavior.[2]

Guidelines for Using DRA

Teachers, therapists, and parents have a long history of using DRA in education, treatment, and everyday social interactions and usually find DRA the easiest of the three differential reinforcement procedures to apply. Practitioners often use DRA as interventions for a wide range of problem behaviors. Many researchers and practitioners have discovered that the following guidelines improve the effectiveness of DRA.

Select Alternative Behavior

Ideally, the alternative behavior (a) already exists in the learner's current repertoire; (b) requires less effort, but never more, than the problem behavior; (c) occurs frequently enough to provide sufficient opportunities for reinforcement; and (d) is likely to receive reinforcement in the learner's natural environment after thinning the DRA schedule. Selecting alternative behaviors meeting these criteria will increase the initial effectiveness of DRA and facilitate the maintenance and generalization of behavior changes after terminating the intervention.

When possible, the practitioner should differentially reinforce alternative behaviors that will lead to or increase the learner's opportunity to acquire useful skills. When no behaviors meeting those criteria can be identified, the practitioner can select an alternative behavior that can be easily taught or consider using a DRO procedure as presented in the next section of this chapter. Figure 25.2 shows some examples of alternative behaviors (Webber & Scheuermann, 1991).

Select Potent Reinforcers and Deliver Them Consistently

Presenting stimulus changes that the practitioner only assumes function as reinforcement can limit or even negate the effectiveness of any reinforcement-based intervention. Providing consequences for the alternative behavior identified by stimulus preference and reinforcer assessments (see Chapter 11) and functional behavior assessments (see Chapter 27) will increase the effectiveness of DRA. In addition, reinforcers should be selected that are relevant to the motivating operations that naturally exist in the treatment setting or that can be created (e.g., by deprivation prior to treatment sessions).

The consequence that maintained the problem behavior prior to intervention is often the most effective reinforcer for the alternative behavior. Results from indirect functional assessments have been used to identify reinforcers for DRA interventions (e.g., Durand, Crimmins, Caufield, & Taylor, 1989; McDonough, Johnson, & Waters, 2017). Functional analyses are most often used to identify maintaining consequences (e.g., Athens & Vollmer, 2010; Dixon et al., 2001; Wilder, Masuda, O'Conner, & Baham, 2001).

The magnitude of the reinforcer used in a differential reinforcement intervention is probably less important than its consistent delivery and control. Lerman, Kelley, Vorndran, Kuhn, and LaRue (2002) provided positive reinforcement (access to

Figure 25.2 Examples of alternative behaviors for common classroom behavior problems.

Problem Behavior	Alternative Behavior
Talking back	Positive response such as "Yes, sir" or "OK" or "I understand"; or acceptable question such as "May I ask you a question about that?" or "May I tell you my side?"
Cursing	Acceptable exclamations such as "darn," "shucks."
Being off-task	Any on-task behavior: looking at book, writing, looking at teacher, etc.
Being out of seat	Sitting in seat (bottom on chair, with body in upright position).
Noncompliance	Following directions within _____ seconds (time limit will depend upon age of student); following directions by second time direction is given.
Talking out	Raising hand and waiting quietly to be called on.
Turning in messy papers	No marks other than answers.
Hitting, pinching, kicking, pushing/shoving	Using verbal expression of anger; pounding fist into hand; sitting or standing next to other students without touching them.
Tardiness	Being in seat when bell rings (or by desired time).
Self-injurious or self-stimulatory behaviors	Sitting with hands on desk or in lap; hands not touching any part of body; head up and not touching anything (desk, shoulder, etc.).
Inappropriate use of materials	Holding/using materials appropriately (e.g., writing *only* on appropriate paper, etc.).

Adapted from "Accentuate the Positive . . . Eliminate the Negative," by J. Webber and B. Scheuermann, 1991, *Teaching Exceptional Children, 24*(1), p. 15. Copyright 1991 by Council for Exceptional Children. Used by permission.

toys) and negative reinforcement (escape from demands) in magnitudes ranging from 20 seconds to 300 seconds. They found that reinforcement magnitude only affected responding minimally during treatment and had little effect on maintenance after treatment was discontinued. However, a series of experiments conducted by Athens and Vollmer (2010) "showed the effectiveness of DRA that provided some combination of more immediate, longer duration, or higher quality of reinforcement for appropriate behavior relative to reinforcement for problem behavior" (p. 586).

Reinforce Alternative Behavior Immediately and Consistently

In addition to their potential effectiveness as reinforcers, the stimulus changes selected by the practitioner must be ones that can be delivered immediately and consistently when the alternative behavior occurs. A continuous schedule of reinforcement (CRF) should be used initially for the alternative behavior and then transition to an intermittent schedule. The reinforcer should be presented consistently and immediately following each occurrence of the alternative behavior. After firmly establishing the alternative behavior, the practitioner should gradually thin the reinforcement schedule.

Withhold Reinforcement for the Problem Behavior

The effectiveness of DRA as an intervention for problem behavior depends on the alternative behavior yielding a higher rate of reinforcement than the problem behavior. Maximizing the difference between rates of reinforcement obtained by the two response classes entails withholding all reinforcement for the problem behavior (i.e., extinction schedule).

Ideally, the alternative behavior would always receive reinforcement (at least initially) and the problem behavior would always be on extinction. As Vollmer, Roane, Ringdahl, and Marcus (1999) noted, "Perfect implementation of differential reinforcement entails providing reinforcers as immediately as possible after an appropriate behavior occurs (e.g., within 5 s). Treatment effects may degrade as the delay to reinforcement increases, especially if inappropriate behavior is occasionally reinforced" (p. 21). However, practitioners must often implement DRA schedules in less than optimal conditions, in which some occurrences of the alternative behavior do not produce reinforcement, and the problem behavior sometimes inadvertently produces the reinforcer.

Results from a study by Vollmer and colleagues (1999) suggest that even when such treatment "mistakes" occur, differential reinforcement may still be effective. These researchers compared a "full implementation" of DRA, in which 100% of instances of alternative behavior were reinforced (CRF) and 0% of instances of aberrant behavior were reinforced (extinction), with various levels of "partial implementation." For example, only one in every four instances of appropriate behavior produced reinforcement with a 25%/75% schedule, whereas a reinforcer followed three of every four instances of inappropriate behavior. As expected, full implementation of differential reinforcement produced the greatest effects, with inappropriate behavior being "virtually replaced by appropriate behavior, and lower levels of implementation eventually reducing treatment efficacy if the schedule of reinforcement favored inappropriate behavior" (p. 20). Figure 25.3 shows the results for Rachel, a 17-year-old student with profound developmental disabilities who engaged in self-injurious behavior (SIB; e.g., head hitting and hand biting) and aggression toward others (e.g., scratching, hitting, hair pulling).

An important and perhaps surprising finding of the Vollmer and colleagues (1999) study was that partial implementation was effective under certain conditions. During partial implementation, the participants' behavior showed a "disproportional tendency toward appropriate behavior" (p. 20) if they had prior exposure to full implementation. The researchers concluded that this finding suggests the possibility of intentionally thinning implementation levels prior to extending differential reinforcement interventions to settings in which maintaining treatment fidelity will be difficult. Recognizing that treatment effects might erode over time, they recommended that full implementation booster sessions be conducted periodically to re-establish the predominance of appropriate behavior.

Watch for Treatment Relapse

A problem behavior previously reduced by DRA may occur more frequently when the reinforcement schedule for the alternative behavior is thinned to levels that can be maintained in the client's everyday environment, when treatment shifts from a clinical setting to the natural settings, or when new therapists and/or caregivers begin working with the client. Problem behavior treated with DRA may ultimately prove more resistant to extinction when the DRA contingency is weakened or compromised (Nevin, Tota, Torquato, & Shull, 1990). Mace and colleagues (2010) conducted basic and translational research experiments exploring the persistence-strengthening effects of DRA and a possible solution.

Resurgence refers to the reoccurrence of a previously reinforced behavior when reinforcement for an alternative behavior is terminated or decreased (Lattal & St. Peter Pipkin, 2009; Lieving & Lattal, 2003). Fuhrman, Fisher, and Greer (2016) described resurgence in the context of FCT "as the recurrence of a previously reinforced response (e.g., destructive behavior) when alternative reinforcement is challenged (e.g., extinction or schedule thinning)" (p. 884). Also, spontaneous recovery (i.e., the reappearance of the problem behavior after it has diminished to its pre-reinforcement level or stopped entirely) and response variation (i.e., the emergence of novel behaviors during the extinction process) relate to resurgence (see Chapter 24). Research has identified several factors that may influence resurgence during applications of DRA, such as magnitude of reinforcement, thinning the schedules of reinforcement, response effort (e.g., Wilson, Glassford, & Koerkenmeier, 2016), and lapses in treatment integrity (e.g., Fuhrman et al., 2016).

The resurgence evoked by extinction might not be the problem behavior; alternative behaviors can resurge as well (Hoffman & Falcomata, 2014). Bloom and Lambert (2015) recommend that clinicians reduce resurgence of the problem behavior by reinforcing multiple alternative behaviors. An analog experiment by Lambert, Bloom, Samaha, Dayton, and Rodewald (2015) provides some empirical support for this

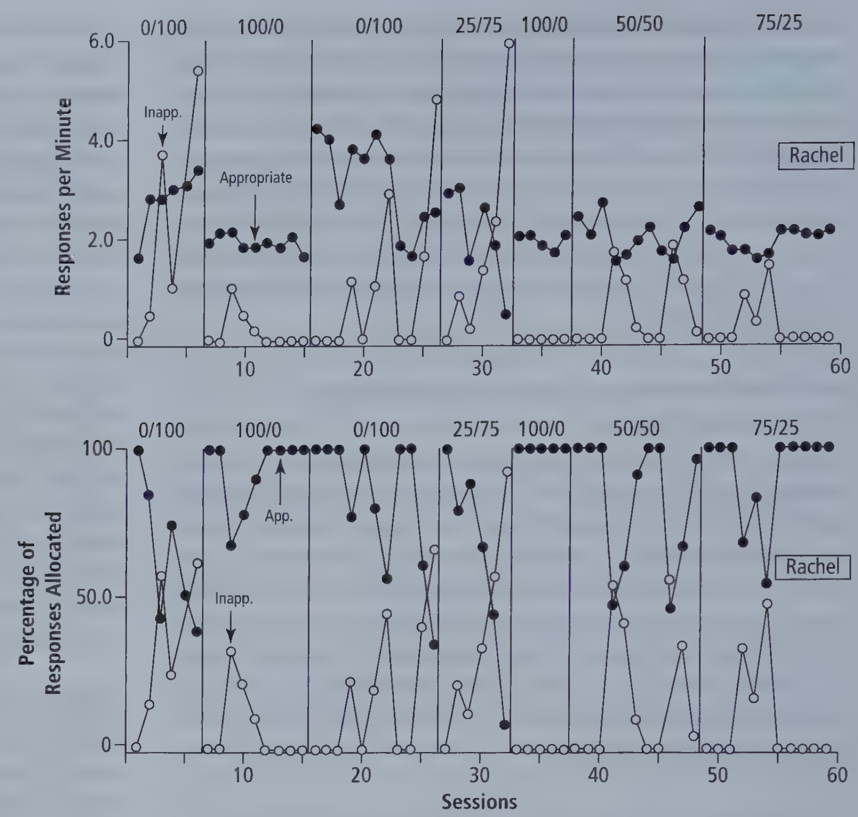


Figure 25.3 Number of appropriate and inappropriate responses per minute (upper panel) and allocation of appropriate and inappropriate behavior (lower panel) by a 17-year-old student with profound developmental disabilities during full (100/0) and partial levels of implementation of DRA.

"Evaluating Treatment Challenges with Differential Reinforcement of Alternative Behavior" by T. R. Vollmer, H. S. Roane, J. E. Ringdahl, and B. A. Marcus, 1999, *Journal of Applied Behavior Analysis*, *32*, p. 17. Reproduced with permission of John Wiley & Sons Inc.

suggestion. Response rates by three women with developmental disabilities on five topographically different devices (i.e., toggle light switch, rocker light switch, slide light switch, cord switch, and personal alarm button) were measured. The researchers arbitrarily identified one switch as the target response (i.e., problem behavior) and all other switches were designated as alternative behaviors.

After an initial *reinforcement phase* of a high-preference edible item on a fixed ratio (FR) 1 schedule developed independent consistent responding on the target response, an *elimination phase* was implemented in which the target response was placed on extinction and an alternative response (or series of three alternative responses) was reinforced on an FR 1 schedule. In a final *resurgence phase*, all responses were placed on extinction. The procedure used in the elimination phase to reinforce one alternative behavior was analogous to traditional DRA intervention. The researchers dubbed the procedure for reinforcing a series of alternative behaviors *serial DRA*. The results are encouraging: "In all cases, the percentage of total responding allocated toward target response resurgence was less in the serial DRA component than in the traditional DRA component . . . Our data provide preliminary evidence suggesting that serial DRA may produce more durable and desirable outcomes than traditional DRA" (p. 765).

Combine DRA with Other Procedures

Considering that DRA interventions do not specifically provide consequences for the problem behavior, a practitioner will seldom apply DRA as a single intervention if the problem behavior is destructive, is dangerous to the learner or to others, or interferes with health and safety. In these situations, the practitioner might combine DRA with other reductive procedures (e.g., response blocking, time-out, stimulus fading, differential reinforcement of other behavior) to produce a more potent intervention (e.g., Goh, Iwata, & Kahng, 1999; Jessel & Borrero, 2014; Patel, Piazza, Kelly, Oschsher, & Santana, 2001).

For example, Ringdahl and colleagues (2002) compared the effects of DRA, with and without instructional fading, on the frequency of problem behaviors of an 8-year-old girl, Kristina, who had been diagnosed with autism and functioned in the moderate range of developmental disabilities. Kristina's problem behaviors during academic tasks—destruction: throwing, tearing, or breaking work materials; aggression: hitting, kicking, biting; and self-injurious behavior: biting herself—had become so severe that she had been admitted to a hospital day treatment program. Because a previous functional analysis revealed that Kristina's problem behavior was maintained by escape from instructions to complete tasks, the DRA procedure in both conditions consisted of giving Kristina a 1-min break from the task and access to leisure materials contingent on independent completion of the task (e.g., counting items, matching cards with the same colors or shapes) without exhibiting problem behavior. Kristina received instructions to complete a work task at approximately one instruction every other minute during the DRA without instructional fading condition. During the DRA with instructional fading, Kristina received no instructions during the first three sessions, then one instruction every 15 minutes. The number of instructions given gradually increased until it equaled the number in the DRA without the fading condition. Problem behavior was high at the outset of the DRA without the instructional fading condition but decreased across sessions to a mean of 0.2 responses per minute over the final three sessions

(see Figure 25.4). During DRA with instructional fading, problem behavior was low from the outset and occurred during only two sessions (Sessions 9 and 14). Over the final three sessions of DRA with instructional fading, Kristina exhibited no problem behavior, and the instruction rate was equal to the DRA without instructional fading (0.5 instruction per minute).

DIFFERENTIAL REINFORCEMENT OF OTHER BEHAVIOR (DRO)

Reynolds (1961) described the differential reinforcement of other behavior as "reinforcement for not responding" (p. 59). A practitioner applying **differential reinforcement of other behavior (DRO)** delivers reinforcement contingent on the problem behavior not occurring throughout intervals of time (*interval DRO*) or at specific moments of time (*momentary DRO*).

Researchers have demonstrated the effectiveness of DRO in reducing a wide range of problem behaviors in a variety of subject populations, including disruptive classroom behavior by typically developing students (Austin, Groves, Reynish, & Francis, 2015), tics in youth with Tourette syndrome (Capriotti et al., 2017), stereotypic vocalizations and pica in children with autism (Rozenblat, Brown, Brown, Reeve, & Reeve, 2009; Slocum et al., 2017), and self-injurious behavior by adults with severe intellectual disabilities (Lindberg, Iwata, Kahng, & DeLeon, 1999). Behavior analysts have used DRO to reduce biting, chewing, and pawing by horses (Fox, Bailey, Hall, & St. Peter, 2012; Fox & Belding, 2015).

Four behavioral processes have been suggested as potential explanations for the effects of DRO (Jessel & Ingvarsson, 2016; Poling & Ryan, 1982).

> First, the presentation of the reinforcer could function as an abolishing operation (i.e., satiation) and temporarily reduce or eliminate target responding. Second, the disruption of the response–reinforcer contingency could result in extinction. Third, delaying access to reinforcers could suppress target responding through punishment. Fourth, other behavior could be strengthened by adventitious reinforcement, resulting in displacement of the target behavior.
>
> It is likely that each of these four processes could be operating on different occasions, or multiple processes could operate at the same time, depending on the DRO arrangement and the reinforcer used. (Jessel & Ingvarsson, 2016, p. 991–992)

Applied behavior analysts implement DRO with two primary methods. With *interval DRO*, if the problem behavior did not occur during a predetermined time interval, a reinforcer is delivered; any instance of the problem behavior during that interval resets the interval and postpones reinforcement.[3] With *momentary DRO*, absence of the problem behavior at predetermined moments in time produces reinforcement; instances of the target behavior at other points in time do not alter when the next opportunity for reinforcement becomes available. The combinations of an interval or momentary omission requirement implemented on either a fixed or a variable schedule yield four basic DRO arrangements shown in Figure 25.5 (Lindberg et al., 1999).

Interval DRO

The following sections introduce procedures for applying fixed-interval DRO, variable-interval DRO, fixed-momentary DRO, and variable-momentary DRO schedules.

Fixed-interval DRO (FI-DRO)

Most applications of interval DRO apply the omission requirement at the end of successive time intervals of equal duration. To apply a **fixed-interval DRO (FI-DRO)** schedule, a behavior analyst (a) establishes a standard interval of time; (b) delivers reinforcement at the end of that interval if the problem did not occur during the interval; and (c) upon any occurrence of the problem behavior, immediately resets the timer to begin a new interval.

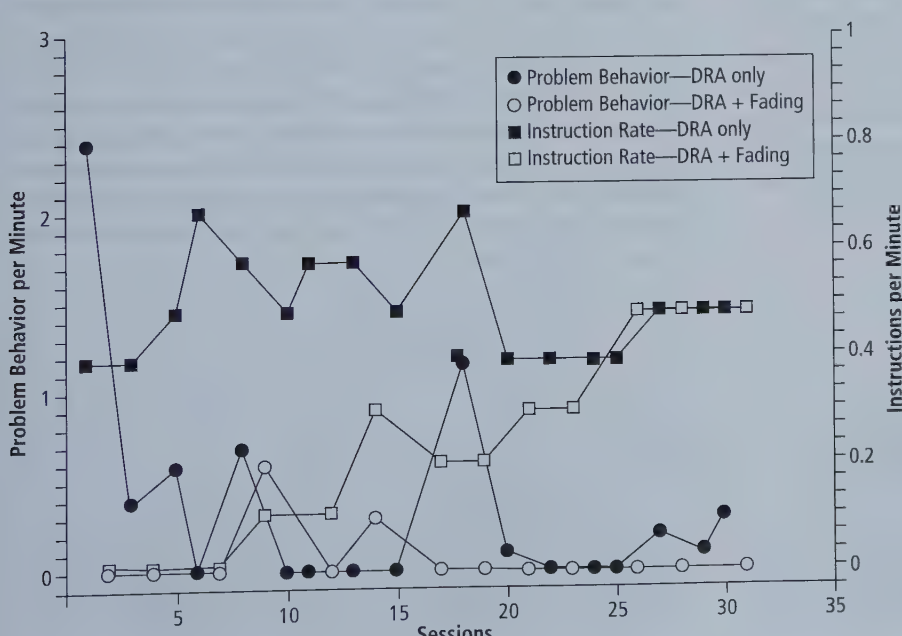


Figure 25.4 Problem behavior by an 8-year-old student with autism and developmental disabilities (left *y*-axis) and instructions given per minute (right *y*-axis) under DRA only and DRA implemented with instructional fading.

Based on "Differential Reinforcement With and Without Instructional Fading" by J. E. Ringdahl, K. Kitsukawa, M. S. Andelman, N. Call, L. C. Winborn, A. Barretto, and G. K. Reed, 2002, *Journal of Applied Behavior Analysis*, *35*, p. 293. Copyright 2002 by the Society for the Experimental Analysis of Behavior, Inc.

For example, Allen, Gottselig, and Boylan (1982) applied an FI-DRO procedure to decrease the disruptive classroom behaviors of third-grade students. They set a kitchen timer to 5-min intervals that continued to run as long as no disruptive behaviors occurred. The timer was reset and a new 5-min interval began if any student engaged in disruptive behavior at any time during the 5-min interval. The class received 1 minute of free time when the timer signaled the end of a 5-min interval without disruptive behavior, which they accumulated and used at the end of the class period.[4]

Behavior analysts can gradually thin the DRO interval and treatment session length as the problem behavior improves (Della Rosa, Fellman, DeBiase, DeQuinzio, & Taylor, 2015). Cowdery and colleagues (1990) used FI-DRO to help Jerry, a 9-year-old who never attended school and spent most of his time hospitalized. He scratched and rubbed his skin so often and hard that he had open lesions all over his body. During treatment, the experimenter left the room and watched Jerry through an observation window. The researchers initially made praise and token reinforcement contingent on 2-min "scratch-free" intervals, which were determined by baseline assessment. If Jerry scratched during the interval, the experimenter entered the room and told him she regretted he had not earned a penny and asked him to try again. Jerry's SIB immediately decreased to zero with the introduction of DRO. Jerry's SIB remained low as the DRO interval and session length were gradually thinned—from three 2-min intervals to three 4-min intervals. Following a brief return to baseline conditions in which Jerry's scratching quickly increased, Cowdery and colleagues reinstated DRO with five intervals per session and increased session length in 1-min increments. Jerry could now earn a total of 10 tokens per session, a token for each interval in which he did not scratch plus a 5-token bonus if he refrained from scratching for all five intervals. Jerry's SIB gradually reduced to zero as session length and DRO intervals were extended. Session 61 consisted of a single 15-min DRO interval, which was followed by one 25- and three 30-min intervals (see Figure 25.6).

Cook, Rapp, and Schulze (2015) used differential negative reinforcement of other behavior (DNRO) with an 8-year-old boy with autism who exhibited problem behavior when a medical alert bracelet was placed on his wrist (e.g., yelling, demands to remove the bracelet, flopping on the floor, hitting others, biting his hand, pulling on and biting the bracelet). For a description of the DNRO procedure and graph of the results, see Chapter 9.

Variable-interval DRO (VI-DRO)

To apply a **variable-interval DRO (VI-DRO)** schedule, a behavior analyst (a) establishes a random sequence of varied intervals of time and (b) delivers reinforcement at the end of each interval if the problem behavior did not occur at any time during the interval. For example, a VI-DRO 10-sec schedule delivers reinforcement contingent on the omission of the target behavior throughout intervals of varying duration that average 10 seconds (e.g., a random sequence of 2-sec, 8-sec, 5-sec, 20-sec, and 15-sec intervals).

Chiang, Iwata, and Dorsey (1979) used VI-DRO to decrease the inappropriate behaviors of a 10-year-old student with developmental disabilities during bus rides to and from school. The boy's disruptive behaviors included aggression (e.g., slapping, poking, hitting, kicking), being out of seat, stereotypic behaviors, and inappropriate vocalizations (e.g., screaming, yelling). Instead of basing the DRO intervals on elapsed time, the researchers used an interesting procedure they described as a "distance-based" DRO schedule. The driver divided the bus route into sections designated by landmarks such as stop signs and traffic lights and mounted a four-digit hand counter on the dashboard of the bus. The counter was within reach of the driver and within view of the student. At each predetermined landmark, the bus driver praised the student's behavior and added one point to the hand counter if no disruptive behaviors had been emitted during the DRO interval. Upon arrival at home or school, the driver recorded the number of points earned on a card and gave the card to the boy's foster father or teacher. The student exchanged the points earned for rewards (e.g., access to toys, home and school privileges, small snacks).

Chiang and colleagues used a two-tier multiple baseline across settings design to evaluate the VI-DRO intervention. During baseline, disruptive behavior ranged from 20% to 100% with an average of 66.2% for the afternoon rides and 0% to 92% with an average of 48.5% in the mornings. When implementing the VI-DRO procedure during the afternoon rides only, disruptive behavior decreased immediately to an average of 5.1% (range of 0% to 40%) during the afternoon rides but remained unchanged during the morning rides. Presenting the intervention also during the morning rides eliminated all disruptive behavior.

Progar and colleagues (2001) implemented VI-DRO as one component of a conjunctive schedule of reinforcement to reduce the aggressive behaviors of Milty, a 14-year-old boy with autism. A conjunctive schedule delivers reinforcement when the requirements of two or more schedule components have been met (see Chapter 13)—in this case, an FR3 schedule for task

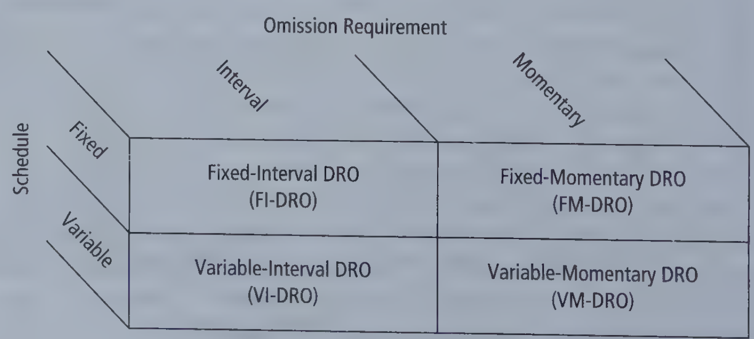


Figure 25.5 Four basic variations of DRO that can be created by combining the omission requirement (interval or momentary) and the schedule of reinforcement (fixed or variable).

"DRO Contingencies: An Analysis of Variable-Momentary Schedules" by J. S. Lindberg, B. A. Iwata, S. W. Kahng, & I. G. DeLeon, 1999. *Journal of Applied Behavior Analysis, 32*, p. 125. Reproduced with permission of John Wiley & Sons Inc.

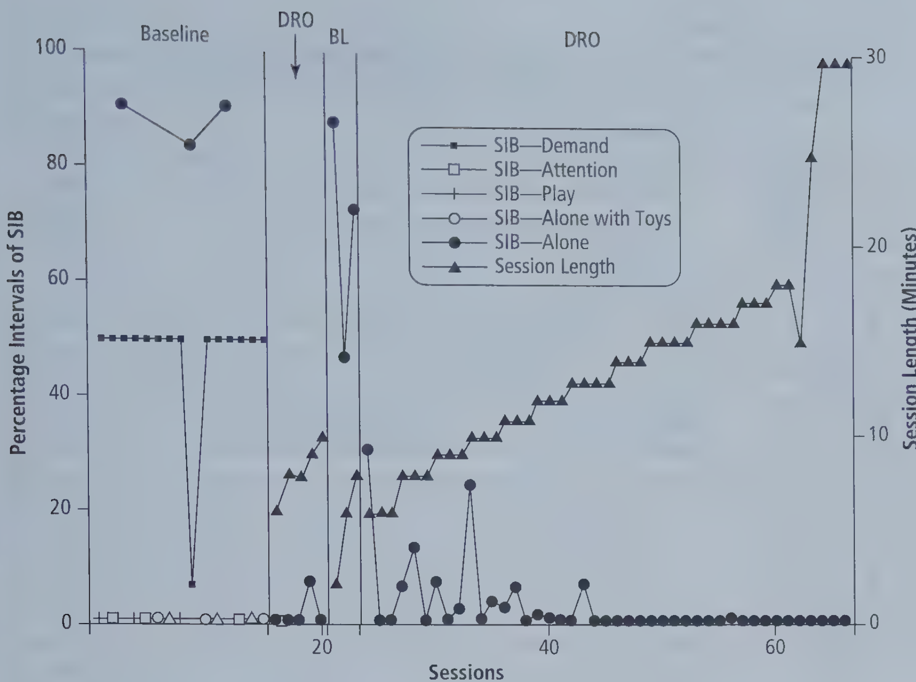


Figure 25.6 Percent intervals of SIB by Jerry during baseline and DRO and duration of DRO sessions (right *y*-axis). Other data points during the first baseline phase show the results of functional analysis, confirming that Jerry's SIB occurred most often when he was alone.

"Effects and Side Effects of DRO as Treatment for Self-injurious Behavior" by G. Cowdery, B. A. Iwata, and G. M. Pace, 1990, *Journal of Applied Behavior Analysis*, *23*, p. 500. Reproduced with permission of John Wiley & Sons Inc.

completion and a VI 148-sec schedule for the absence of aggression. Milty received an edible reinforcer each time he completed three components of a task (e.g., making his bed, vacuuming, straightening objects in his room) in the absence of aggression for the duration of the variable interval. An occurrence of aggression prior to completing the FR 3 requirement reset both components of the conjunctive schedule. This intervention produced a substantial reduction of aggressive behavior.

Momentary DRO

On a momentary DRO schedule, reinforcement is contingent on the absence of the problem behavior at the exact time each interval ends, rather than throughout the entire interval, as with interval DRO. Occurrences of the problem behavior between those moments have no effect.

Fixed-momentary DRO (FM-DRO)

To implement a **fixed-momentary DRO (FM-DRO)** schedule, the practitioner (a) establishes a standard interval of time; (b) observes the participant at the moment each interval ends; and (b) delivers reinforcement at the end of each interval if the problem behavior is not occurring at that moment.

In their review of research on momentary DRO schedules, Hammond, Iwata, Fritz, and Dempsey (2011) concluded that FM-DRO schedules "generally have been ineffective as treatment for problem behavior. Because most early research on FM DRO included presentation of a signal at the end of the DRO interval, it is unclear whether the limited effects of FM-DRO were due to (a) the momentary response requirement of the schedule per se or (b) discrimination of the contingency made more salient by the signal" (p. 69).[5] The researchers addressed this question by comparing the effects of signaled versus unsignaled FM-DRO. Four students with special needs served as participants—(1) Seth, a deaf and learning-impaired, aggressive 6-year-old; (2) Alex, with autism and seizure disorder, an aggressive 14-year-old; (3) Curtis, with autism, an aggressive 13-year-old; and (4) Abby, a 19-year-old with profound intellectual disabilities, autism, and seizure disorder. Functional analyses revealed that positive reinforcement in the form of access to preferred items maintained the participants' problem behaviors.

The four participants had access to preferred items during baseline. The signaled and unsignaled conditions were identical except for the signal.

> *Signaled FM DRO.* The experimenter covertly held the reinforcer to ensure that it could be delivered quickly but that no additional movements (e.g., reaching for the reinforcer) might inadvertently signal the end of an interval. Three seconds prior to the end of an interval, the experimenter (positioned directly in front of or next to the subject) signaled impending reinforcement delivery by holding the reinforcer above her head. If no problem behavior was occurring at the precise second the interval clocked out, the reinforcer (30-s access to the toy box or video, or one edible item) was delivered. If problem behavior was occurring at that second, the signal was removed, and the reinforcer was withheld until the next scheduled interval.
>
> *Unsignaled FM DRO.* This condition was similar to the signaled condition, except that the experimenter did not signal impending reinforcer delivery. Rather, the experimenter quickly delivered the reinforcer contingent on the absence of problem behavior when the interval ended. (p. 75)

Seth and Alex had rapid reductions in aggression during the signaled FM-DRO conditions (see Figure 25.7). However, the signaled FM-DRO conditions produced different results for Curtis and Abby (see Figure 25.8). Curtis's aggression increased

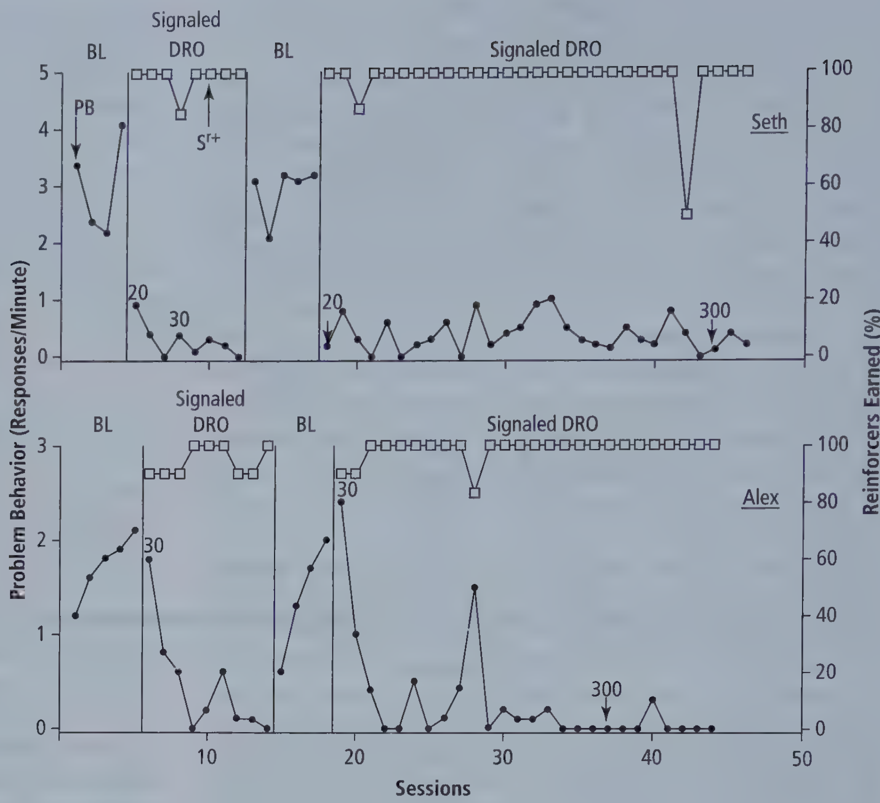


Figure 25.7 Responses per minute of problem behavior (PB) during baseline (BL) and signaled DRO conditions, and percentage of available reinforcers (S^{r+}) earned during DRO conditions for Seth and Alex. Numbers above the data path for problem behavior indicate DRO interval lengths (in seconds) that were in effect during the first and final sessions of each phase.

"Evaluation of Fixed Momentary DRO Schedules Under Signaled and Unsignaled Arrangements" by J. L. Hammond, B. A. Iwata, J. N. Fritz, and C. M. Dempsey, 2011, *Journal of Applied Behavior Analysis*, *44*, 76. Reproduced with permission of John Wiley & Sons Inc.

during the first introduction of the signaled FM-DRO condition, then greatly decreased by the end of the first introduction. His aggression did not decrease during the second signaled FM-DRO condition, but decreased to zero during the reintroduction of the unsignaled FM-DRO condition. Similar to Curtis's behavior, Abby's head hitting decreased under the second and final unsignaled DRO phase, all the while earning a large proportion of available reinforcers. "[T]hese data indicate that momentary DRO should be used with caution, because very subtle therapist responses (e.g., reaching for the programmed reinforcer) may function as signals that could render FM DRO ineffective" (Hammond et al., 2011, p. 80).

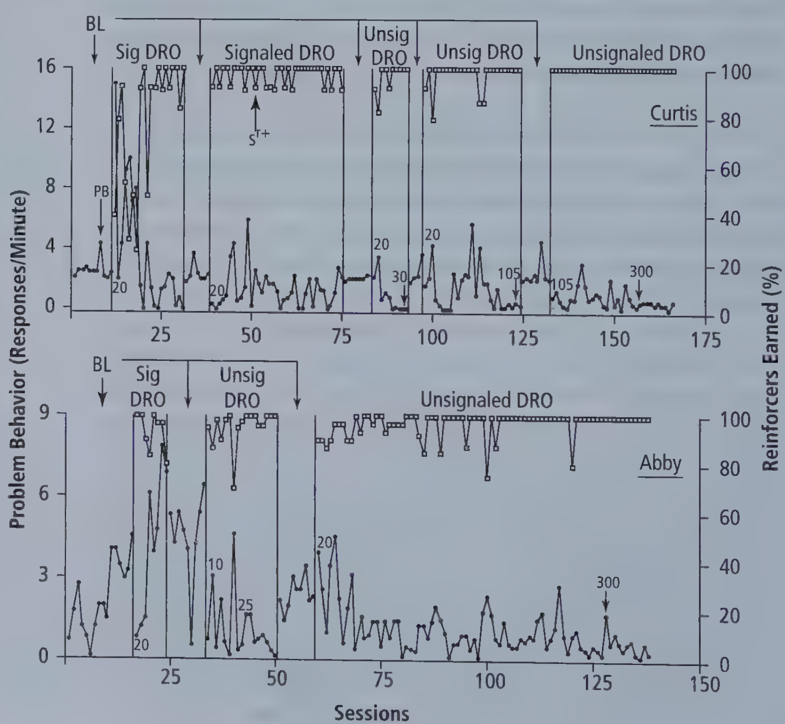


Figure 25.8 Responses per minute of problem behavior (PB) during baseline (BL) and signaled DRO conditions, and percentage of available reinforcers (S^{r+}) earned during DRO conditions for Curtis and Abby. Numbers above the data path for problem behavior indicate DRO interval lengths (in seconds) that were in effect during the first and final sessions of each phase.

"Evaluation of Fixed Momentary DRO Schedules Under Signaled and Unsignaled Arrangements" by J. L. Hammond, B. A. Iwata, J. N. Fritz, and C. M. Dempsey, 2011, *Journal of Applied Behavior Analysis*, *44*, p. 76. Reproduced with permission of John Wiley & Sons Inc.

Variable-momentary DRO (VM-DRO)

In a **variable-momentary DRO (VM-DRO)** schedule, the practitioner (a) establishes a random sequence of varied intervals of time and (b) delivers reinforcement at the end of each interval if the problem behavior is not occurring at that precise moment. Lindberg and colleagues (1999) used a VM-DRO intervention to help Bridget, a 50-year-old woman with severe developmental disabilities, who often banged her head and body. A functional analysis indicated that social-positive reinforcement maintained Bridget's self-injurious behaviors (SIB). During the VM-DRO intervention, SIB was placed on extinction and Bridget received 3 to 5 seconds of attention from the therapist if she did not bang her head or hit herself when each interval ended. A VM-DRO 15-sec schedule was implemented for five sessions following baseline, and Bridget's SIB decreased abruptly to almost zero (see Figure 25.9). Following a second baseline phase in which Bridget's SIB increased, the researchers reintroduced VM-DRO with a mean interval of 11 seconds. SIB did not decrease as quickly during the second treatment condition. After several sessions in which SIB occurred at low rates, the interval length was increased to 22 seconds and eventually to the maximum target interval of 300 seconds.

Interval DRO has been used more widely than momentary DRO. Some researchers have found that interval DRO is more effective than momentary DRO for suppressing problem behavior, and that momentary DRO might be most useful in maintaining reduced levels of problem behavior produced by interval DRO (Barton, Brulle, & Repp, 1986; Repp, Barton, & Brulle, 1983). Lindberg and colleagues (1999) noted two potential advantages of VM-DRO schedules over FI-DRO schedules. First, the VM-DRO schedule appears more practical because the practitioner need not monitor the participant's behavior at all times. Second, data obtained by these researchers showed that participants obtained higher overall rates of reinforcement with VM-DRO than with FI-DRO.

Guidelines for Using DRO

In addition to the importance of selecting potent reinforcers, we recommend the following guidelines for the effective use of DRO.

Set Initial DRO Intervals That Ensure Frequent Reinforcement

Practitioners should establish an initial DRO time interval that ensures that the learner's current level of behavior will contact reinforcement when the DRO contingency is applied. For example, Cowdery and colleagues (1990) initially made praise and token reinforcement contingent on 2-min "scratch-free" intervals because that was "the longest we had previously seen Jerry refrain from scratching while left alone" (p. 501). Beginning with an interval that is equal to or slightly less than the mean baseline interresponse time (IRT) usually will make an effective initial DRO interval. To calculate a mean IRT, the practitioner divides the total duration of all baseline measurements by the total number of responses recorded during baseline. For example, if a total of 90 responses were recorded during 30 minutes of baseline measurement, the mean IRT would be 20 seconds (i.e., 1800 seconds ÷ 90 responses). Using these baseline data as a guide, the practitioner could set the initial DRO interval at 30 seconds or less.

Do Not Inadvertently Reinforce Other Undesirable Behaviors

Reinforcement on a "pure" DRO schedule is contingent on a very general response class that is defined only by the absence of the targeted problem behavior. Consequently, a "pure" DRO is typically used as a treatment of choice for serious behavior problems that occur at very high rates by people whose current repertoires provide few, if any, other behaviors that might function as alternative behaviors and for whom just about anything else they might do is less of a problem than the target behavior.

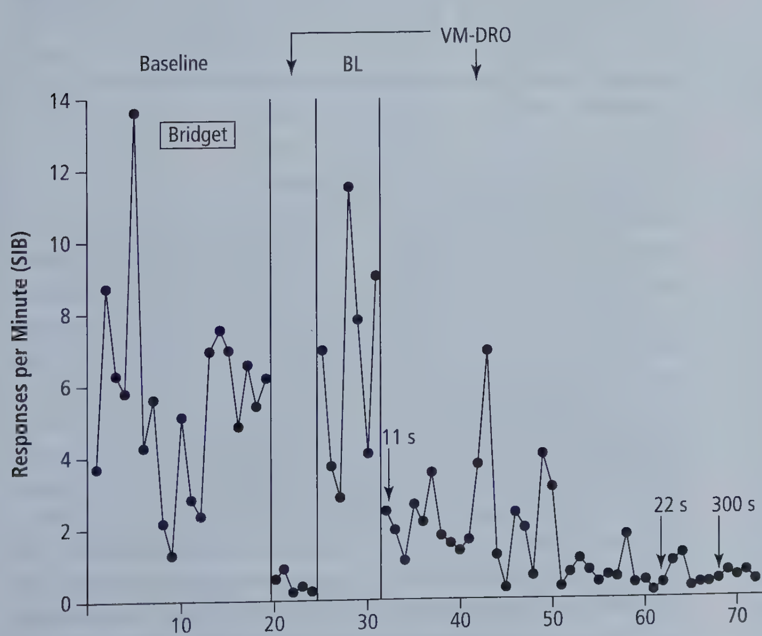


Figure 25.9 Responses per minute of self-injurious behavior (SIB) during baseline and treatment conditions.

"DRO Contingencies: An Analysis of Variable-momentary Schedules" by J. S. Lindberg, B. A. Iwata, S. W. Kahng, and I. G. DeLeon, 1999. *Journal of Applied Behavior Analysis, 32*, p. 131. Reproduced with permission of John Wiley & Sons Inc.

Because DRO does not require any certain behaviors to be emitted for reinforcement, whatever the person is doing when reinforcement is delivered is likely to occur more often in the future (Jessel, Borrero, & Becraft, 2015). Therefore, practitioners using DRO must be careful not to strengthen inadvertently other undesirable behaviors. When using DRO, the practitioner should deliver reinforcement at the intervals or moments in time specified by the schedule contingent on the absence of the problem behavior *and* the absence of any other significant inappropriate behaviors.

Gradually Increase the DRO Interval

After the initial DRO interval effectively controls the problem behavior, the practitioner can increase the DRO interval, by increasing the interval through a series of initially small and gradually increasing increments.

Poling and Ryan (1982) suggested three procedures for increasing the duration of the DRO interval.

1. Increase the DRO interval by a *constant duration* of time. For example, a practitioner could, at each opportunity for an interval increase, lengthen the DRO interval by 15 seconds.
2. Increase intervals *proportionately*. For example, a practitioner, at each opportunity to change the interval, could increase the DRO interval by 10%.
3. Change the DRO interval each session based on the *learner's performance*. For example, a practitioner could set the DRO interval for each session equal to the mean IRT from the preceding session.

If problem behavior worsens when a longer DRO interval is introduced, the practitioner can decrease the duration of the interval to a level that again controls the problem behavior. The DRO interval can then be extended with smaller, more gradual durations after previously obtained reductions in problem behavior have been re-established.

Extend the Application of DRO to Other Settings and Times of Day

When the frequency of the problem behavior is reduced substantially in the treatment setting, the DRO intervention can be introduced during other activities and times in the person's natural environment. Having teachers, parents, or other caregivers begin to deliver reinforcement on the DRO schedule will help extend the effects. For example, Cowdery and colleagues (1990) began to implement the DRO intervention at other times and places during the day after Jerry's SIB was under good control during treatment sessions (see Figure 25.6). At first, various staff members would praise Jerry and give him tokens for refraining from scratching during single 30-min DRO intervals during leisure activities, instructional sessions, or free time. Then they began adding extra 30-min DRO intervals until the DRO contingency was extended to all of Jerry's waking hours. Jerry's SIB decreased sufficiently to allow him to be discharged to his home for the first time in 2 years. Jerry's parents were taught how to implement the DRO procedure at home.

Combine DRO with Other Procedures

DRO can be used as a single intervention. However, as with DRA, including DRO in a treatment package with other behavior-reduction procedures often yields more efficient and effective behavior change. De Zubicaray and Clair (1998) used a combination of DRO, DRI, and a restitution (punishment) procedure to reduce the physical and verbal abuse and physical aggression of a 46-year-old woman with intellectual disabilities who lived in a residential facility for people with chronic psychiatric disorders. Further, the results of three experiments by Rolider and Van Houten (1984) showed that an intervention consisting of DRO plus reprimands was more effective than DRO alone in decreasing a variety of problem behaviors by children (e.g., physical abuse by a 4-year-old girl toward her baby sister, thumb sucking by a 12-year-old boy, and bedtime tantrums).

DRO can also be added as a supplement to an intervention that has produced insufficient results. McCord, Iwata, Galensky, Ellingson, and Thomson (2001) added DRO to a stimulus fading intervention that was "having limited success" in decreasing the problem behaviors of Sarah, a 41-year-old woman with severe intellectual disabilities and visual impairment. Her problem behaviors included severe self-injurious behavior, property destruction, and aggression evoked by noise. Gradually increasing the noise volume by 2 dB when DRO was in effect, "a therapist delivered a preferred edible item (half a cheese puff) to Sarah following each 6-s interval in which problem behavior was not observed. If problem behavior occurred, the food was withheld and the DRO interval was reset" (p. 455). Each time Sarah completed three consecutive 1-min sessions with no occurrences of the problem behavior, the researchers increased the DRO interval by 2 seconds. This schedule thinning continued until Sarah could complete the session with no occurrences of the problem behavior. The therapist then delivered the edible item only at the end of the session. The addition of DRO to the intervention produced an immediate decrease in Sarah's problem behavior, which remained at or near zero in the treatment setting for more than 40 sessions (see Chapter 6, the bottom graph of Figure 6.6). Probes conducted in noisy conditions after treatment ended showed that the treatment effects had maintained and generalized to Sarah's home.

Recognize the Limitations of DRO

Although DRO can be highly effective for reducing problem behaviors, it is not without shortcomings. With interval DRO, reinforcement is delivered contingent only on the absence of problem behavior during the interval, even though another inappropriate behavior might have occurred during that time. For instance, let us assume that a 20-sec interval DRO is to be implemented to reduce the facial tics of an adolescent with Tourette syndrome. Reinforcement will be delivered at the end of each 20-sec interval containing no facial tics. However, if the adolescent engages in cursing behavior at any time during the interval or at the end of the interval, reinforcement will still be delivered. It is possible that reinforcement delivered contingent on the absence of facial tics will occur in close temporal proximity to cursing, thereby inadvertently strengthening

another inappropriate behavior. In such cases, the length of the DRO interval should be shortened and/or the definition of the problem behavior expanded to include the other undesirable behaviors (e.g., reinforcement contingent on the absence of facial tics *and* cursing).

With momentary DRO, reinforcement is delivered contingent on the problem behavior not occurring at the end of each interval, even though the inappropriate behavior might have occurred during the interval. Continuing the previous example, on an FM-DRO 20-sec schedule, reinforcement is delivered at each successive 20th second if facial tics are not occurring at those precise moments, even if tics occurred for 50%, 75%, or even 95% of the interval. In such circumstances, the practitioner should use interval DRO and reduce the length of the DRO interval.

Another consideration is the number of resources that may be available for labor-intensive procedures. Rozenblat and colleagues (2009) found that if DRO intervals are extremely short—in their case, as low as 1 sec—it might be clinically impossible for a single analyst or practitioner to implement the procedure. Working alone and trying to conduct a session with other students, time a series of intervals, collect data and record performance, and deliver reinforcers is likely to exceed a single person's capability. Thus, practitioners would be advised to assess staff resources before launching a DRO program.

DIFFERENTIAL REINFORCEMENT OF LOW RATES OF RESPONDING (DRL)

A practitioner applying **differential reinforcement of low rates (DRL)** delivers reinforcement (a) following each occurrence of the target behavior that is separated from the previous response by a minimum amount of time or (b) contingent on the number of responses within a period not exceeding a predetermined criterion. In laboratory research, Ferster and Skinner (1957) found that delivering a reinforcer following a response that had been preceded by increasingly longer intervals of time without a response reduced the overall rate of responding. Catania (2013) describes typical DRL behavior found in laboratory research.

> In DRL performance, responding is unlikely to extinguish, because decreasing low rates produce even more reinforcers. Once some responding has been reinforced, response rate typically increases, but once it has done so responses follow each other more closely so fewer meet the scheduled criterion, and response rate decreases once more, and so on. Typically the responding stabilizes at some value, oscillating between increased rates accompanied by decreased reinforcement and decreased rates accompanied by increased reinforcement. A pigeon's pecking can be maintained over long periods by such contingencies. Thus, DRL performance is highly resistant to change, though its rate remains relatively low. It therefore again illustrates the independence of response strength and response rate (Nevin & Grace, 2000). (p. 209)

Because reinforcement with a DRL schedule is delivered following an *occurrence* of the target behavior—as contrasted with DRO, in which reinforcement is contingent on the *absence* of the behavior—behavior analysts use DRL to decrease the rate of a behavior that occurs too frequently, but not to eliminate the behavior entirely. A practitioner may identify a certain behavior as a problem behavior not because of its form, but because the behavior occurs too often. For example, a student asking for help by raising his hand or coming to the teacher's desk during independent seatwork is appropriate at a limited rate, but a problem behavior if it occurs too frequently.

Some learners may emit a behavior less frequently when told to do the behavior less often, or when given rules stating an appropriate rate of response. When instruction alone does not reduce the behavior to an acceptable rate, the practitioner may need to use a consequence-based intervention. DRL offers practitioners an intervention for diminishing problem behaviors that is stronger than instruction alone but still falls short of more restrictive consequences (e.g., punishment). Deitz (1977) named and described three DRL procedures: full-session DRL, interval DRL, and spaced-responding DRL.

Full-session DRL

In a **full-session DRL** schedule, reinforcement is delivered when responding during an entire instructional or treatment session is equal to or below a predetermined criterion. If the number of responses exceeds the specified limit during the session, reinforcement is withheld. For example, a teacher applying full-session DRL with a criterion limit of four disruptions per class period would provide reinforcement at the end of the class period contingent on four or fewer disruptions. Full-session DRL is an easy-to-apply intervention for problem behaviors in education and treatment settings.

Deitz and Repp (1973) demonstrated the efficacy and manageability of full-session DRL for diminishing classroom misbehavior. They decreased the talk-outs of an 11-year-old boy with developmental disabilities. During 10 days of baseline, the student averaged 5.7 talk-outs per 50-minute session. With the introduction of full-session DRL, the student was allowed 5 minutes for play at the end of the day contingent on 3 or fewer talk-outs during a 50-minute session. Talk-outs during the DRL condition decreased to an average of 0.93 talk-out per 50 minutes. A return to baseline slightly increased the talk-outs to an average of 1.5 per session.

Austin and Bevan (2011) implemented a variation of full-session DRL with three female elementary students who engaged in excessive attention-seeking behaviors that affected work completion in a timely manner. The researchers set the number of requests for assistance each student was allowed to "reflect not only a reduction in baseline rates but also the number of requests the teacher thought was reasonable for each child to make in a 20-min work period" (p. 454). The DRL intervention was implemented as follows:

> At the beginning of independent work sessions, each child received a small index card with boxes corresponding to the number of requests allowed during the session plus one (e.g., Elin was allowed three bids for attention, so her card had four boxes). Each time the child requested attention from the teacher, the teacher responded to the child's request (up to the number of responses allocated for the

session) and initialed one of the boxes on the child's card. If the child exceeded the number of allowable requests, the teacher was instructed to initial the remaining box but not to interact with the child. However, none of the children ever exceeded the number of responses eligible for attention during the DRL conditions. If the child had at least one blank box at the end of the work session, she earned a point for her team. No additional points were awarded if more than one box was left empty, because we wanted the children to use the opportunities for assistance or attention that were available to them. (p. 454)

Because the children's requests within the DRL-imposed limit were reinforced as they occurred during the session, an appropriate procedure given the target behavior, we identify Austin and Bevan's (2011) intervention as a variation of full-session DRL. A reversal design demonstrated the intervention's effectiveness in reducing each student's rate of attention-seeking behavior (see Figure 25.10). The procedure earned high marks on social validity assessments conducted at the end of the study. The teacher found the DRL schedule easy to implement and integrate into ongoing classroom routines and reported she planned to continue using it. The children indicated they were happy to use the check-box cards, liked earning points for their group, and wanted their teacher to continue using the cards.

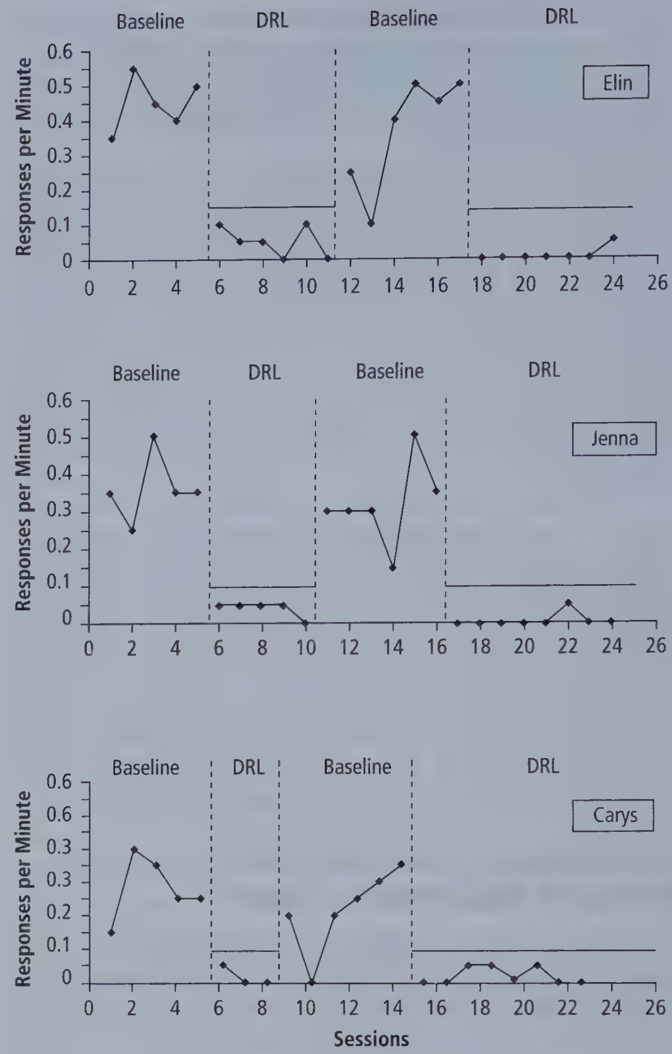


Figure 25.10 Rates of requesting attention across conditions. The solid horizontal line in the DRL phases represents the limit on responding specified by the DRL schedule.

"Using Differential Reinforcement of Low Rates to Reduce Children's Requests for Teacher Attention" by J. L. Austin and D. Bevan, 2011, *Journal of Applied Behavior Analysis*, *44*, p. 456. Used by permission. Reproduced with permission of John Wiley & Sons Inc.

Interval DRL

To apply an **interval DRL** schedule of reinforcement, the practitioner divides the total session into a series of equal intervals of time and provides reinforcement at the end of each interval in which the number of occurrences of the problem behavior was equal to or below a criterion limit. If the learner exceeds the criterion number of responses during an interval, the practitioner removes the opportunity for reinforcement and begins a new interval.

Deitz and colleagues (1978) used an interval DRL schedule of reinforcement to reduce the disruptive behaviors of a student with learning disabilities. The 7-year-old student had several difficult classroom misbehaviors (e.g., running, shoving, pushing, hitting, throwing objects). The student received a sheet of paper ruled into 15 blocks, with each block representing a 2-min interval. A star was placed in a block each time the student completed 2 minutes with 1 or no misbehaviors. Each star permitted the student to spend 1 minute on the playground with the teacher. If the student emitted 2 misbehaviors during the interval, the teacher immediately started a new 2-min interval.

Spaced-responding DRL

Using a **spaced-responding DRL** schedule of reinforcement, the practitioner delivers a reinforcer following an occurrence of a response that is separated from the previous response by at least a specified amount of time.[6] As you will recall from Chapter 4, *interresponse time (IRT)* is the technical term for the duration of time between two responses. IRT and rate of responding are directly correlated: The longer the IRT, the lower the overall rate of responding; shorter IRTs correlate with higher response rates. When reinforcement is contingent on longer IRTs, response rate will decrease.

Favell, McGimsey, and Jones (1980) used a spaced-responding DRL schedule of reinforcement and response prompts to decrease the rapid eating of four persons with profound developmental disabilities. At the start of treatment, reinforcement was contingent on short independent pauses (IRTs) between bites of food. As treatment progressed, gradually increasing longer and longer pauses between bites were required for reinforcement. The researchers also manually prompted a separation between bites of food and faded out those response prompts when a minimum of 5-sec pauses occurred independently between approximately 75% of all bites. Finally, Favell and colleagues gradually thinned food reinforcement and praise. The frequency of eating was decreased from a pretreatment baseline average of 10 to 12 bites per 30 seconds to 3 to 4 bites per 30 seconds during the spaced-responding DRL condition.

Singh, Dawson, and Manning (1981) used a spaced-responding DRL intervention to reduce the stereotypic responding (e.g., repetitive body movements, rocking) of three teenage girls with profound developmental disabilities. During the first phase of spaced-responding DRL intervention, the therapist praised each girl whenever she emitted a stereotypic response at least 12 seconds after the last response. One of the experimenters timed the intervals and used a system of automated lights to signal the therapist when reinforcement was available. After the DRL 12-sec IRT resulted in an abrupt decrease in stereotypic behavior by all three girls (see Figure 25.11), Singh and colleagues systematically increased the IRT criterion to 30, 60, and then 180 sec. The spaced-responding DRL procedure not only produced substantial reductions in stereotypic responding in all three subjects but also had the concomitant effect of increasing appropriate behavior (e.g., smiling, talking, playing).

Most behavior-reduction procedures that manipulate consequences have the potential to reduce behavior to zero occurrences. Spaced-responding DRL is unlikely to do that. For that reason, spaced-responding DRL is an important intervention for diminishing behavior that should still continue but at a slower rate. A spaced-responding DRL contingency tells learners that their behavior is acceptable, but they should do it less often. For example, a teacher could use a spaced-responding DRL schedule to diminish a student's problem behavior of asking too many questions. The questions occurred so frequently that they interfered with the classroom learning and teaching. To intervene, the teacher could respond to the student's question if

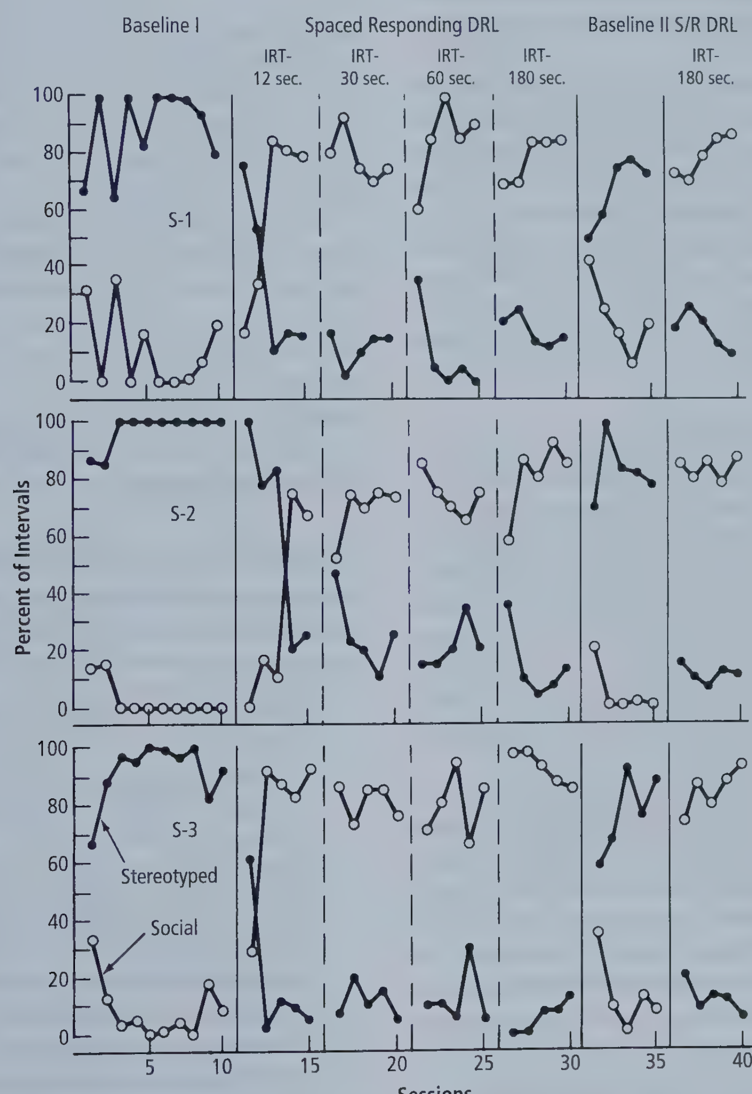


Figure 25.11 Effects of spaced-responding DRL on stereotypic responding of three teenage girls with profound developmental disabilities.

"Effects of Spaced Responding DRL on the Stereotyped Behavior of Profoundly Retarded Persons" by N. N. Singh, M. J. Dawson, and P. Manning, 1981. *Journal of Applied Behavior Analysis, 38*, p. 524. Reproduced with permission of John Wiley & Sons Inc.

the student had not asked a question during a preceding minimum of 5 minutes. This spaced-responding DRL intervention could decrease, yet not eliminate, asking questions.

Guidelines for Using DRL

Several factors influence the effectiveness of the three DRL schedules in diminishing problem behaviors. The following guidelines address those factors.

Choose the Most Appropriate DRL Procedure

Full-session, interval, and spaced-responding DRL schedules provide different levels of reinforcement for learners. Of the three DRL procedures, only spaced-responding DRL delivers reinforcement immediately following the occurrence of specific response; a response must be emitted following a minimum IRT before reinforcement. Practitioners use spaced-responding DRL to reduce the occurrences of a behavior while maintaining those behaviors at lower rates.

With full-session and interval DRL, a response does not need to occur for the participant to receive the reinforcer. Practitioners may apply full-session or interval DRL when it is acceptable that the rate of the problem behavior reaches zero or as an initial step toward the goal of eliminating the behavior.

Spaced-responding and interval DRL usually produce reinforcement at a higher rate than full-session DRL does. Arranging frequent contact with the reinforcement contingency is especially appropriate, and most often necessary, for learners with severe problem behaviors.

Recognize the Limitations of DRL

If a practitioner needs to reduce an inappropriate behavior quickly, DRL would not be the method of first choice. Reducing the inappropriate behavior to appropriate levels with DRL may take more time than the practitioner can afford. Further, DRL should not be used with self-injurious, violent, or potentially dangerous behaviors. Finally, from a practical standpoint, using DRL means that the practitioner must focus on the inappropriate behavior. If a teacher, for example, is not cautious, he may unintentionally give too much attention to the inappropriate behavior, inadvertently reinforcing it.

Effective application of interval DRL and spaced-responding DRL procedures requires continuous monitoring of problem behavior, careful timing, and frequent reinforcement. Without the help of an assistant, practitioners may have difficulty applying the interval DRL procedure in group settings. Interval and spaced-responding DRL procedures are quite reasonable for one-on-one instruction or when competent assistance is available. Jessel and Borrero (2014) offer guidance for the selection and application of DRL schedules.

> Evaluation and implementation of spaced-responding DRL and full-session DRL should be considered in the context of clinical or research goals. The more effortful approach (spaced responding DRL) may be required when the goal is to sustain responding, albeit at rates lower than those produced under baseline conditions (e.g., applications to rapid eating, excessive hand raising, or tattling). In addition, the limitations of using spaced-responding DRL schedules (e.g., requires individual timers for each student to be reset following responding below the programmed IRT) may be assuaged when considering the possibility of the use of widespread technology such as tablets in classroom settings. Applications could be created on tablets with which teachers could select the IRT for each student immediately before the start of classroom activities. The teacher would only have to click on the child's name after each response to determine whether or not he or she met the criteria for reinforcer presentation. (p. 321)

Use Baseline Data to Guide the Selection of the Initial Response or IRT Limits

Practitioners can use the mean number of responses emitted during baseline sessions, or slightly lower than that average, as the initial full-session DRL criterion. For example, 8, 13, 10, 7, and 12 responses recorded per session over five baseline sessions equals a mean of 10 responses per session. Therefore, a limit of 8 to 10 responses per session would be an appropriate initial full-session DRL criterion.

Interval DRL and spaced-responding DRL initial time criteria can be set at the baseline mean or slightly lower. For example, 1 response per 15 minutes makes an acceptable initial interval DRL criterion, as calculated from a baseline mean of 4 responses per 60-min session. With the same baseline data (i.e., 4 responses per 60-min session), it appears reasonable to use 15 minutes as the initial IRT criterion for a spaced-responding DRL schedule. That is, a response will produce reinforcement only if it is separated from the prior response by a minimum of 15 minutes.

Gradually Thin the DRL Schedule

Practitioners should gradually thin the DRL schedule to achieve the desired final rate of responding. Practitioners commonly use three procedures for thinning the initial DRL time criterion.

1. *With full-session DRL*, the practitioner can set a new DRL criterion using the participant's current DRL performances. Another option is to set a new DRL criterion at slightly less than the mean number of responses emitted during recent DRL sessions.
2. *With interval DRL*, the practitioner can gradually decrease the number of responses per interval if the current criterion is more than one response per interval; or, gradually increase the duration of the criterion interval if the current criterion is one response per interval.
3. *With spaced-responding DRL*, the practitioner can adjust the IRT criterion based on the mean IRT of recent sessions, or slightly less than that average. For example, Wright and Vollmer (2002) set the IRT for the DRL component of a treatment package that successfully reduced the rapid eating of a 17-year-old girl at the mean IRT of the previous five sessions. The researchers did not exceed an IRT of 15 seconds because it was not necessary to reduce the girl's eating below a rate of four bites per minute.

Practitioners who successfully thin DRL schedules of reinforcement make gradual, but systematic, changes in time and response criteria associated with full, interval, and space-responding DRL variations.

Two possible decision rules for thinning the DRL schedule are as follows:

> ***Rule 1:*** Practitioners may want to change the DRL criterion whenever the learner meets or exceeds the criterion for three consecutive sessions.
>
> ***Rule 2:*** Practitioners may want to change the DRL criterion whenever the learner receives reinforcement for at least 90% of the opportunities for three consecutive sessions.

Provide Feedback

The effectiveness of a DRL procedure can be enhanced by feedback to help the learner monitor her rate of responding. Full-session, interval, and spaced-responding DRL procedures provide different levels of feedback for participants. The most accurate feedback comes with spaced-responding DRL because reinforcement follows each response that meets the IRT criterion. When a response occurs that does not meet the IRT criterion, reinforcement is withheld, the time interval is immediately reset, and a new interval begins. This process provides learners with immediate feedback on responding per interval.

Interval DRL also provides a high level of feedback, although it is less than that of spaced-responding DRL. Interval DRL provides the learner with two types of feedback. The first problem behavior does not provide feedback. The second response, however, resets the time interval, providing a consequence for the problem behavior. Reinforcement occurs at the end of the interval when one or no problem behaviors occurred during the interval. These two types of feedback improve the effectiveness of the interval DRL intervention (Deitz et al., 1978).

Applied behavior analysts can arrange full-session DRL with or without feedback. The usual arrangement does not provide feedback concerning moment-to-moment accumulation of responses. Deitz (1977) stated that with full-session DRL, learners would respond only to the DRL criterion: Once the learner loses the opportunity for reinforcement during the session, he could then emit high rates of misbehavior without consequence. When the schedule is arranged without moment-to-moment feedback, learners usually stay well below the DRL limit. Full-session DRL, without moment-to-moment feedback, may not be as effective for learners with severe problem behaviors as spaced-responding and interval DRL. The effectiveness of full-session DRL relies heavily on an initial verbal description of the contingencies of reinforcement (Deitz, 1977).

SUMMARY

Differential Reinforcement Defined

1. Differential reinforcement entails reinforcing one response class and withholding reinforcement for another response class.

2. When used as a reductive procedure for problem behavior, differential reinforcement consists of (a) providing reinforcement contingent on a behavior other than the problem behavior (or the problem behavior occurring at a reduced rate) and (b) withholding reinforcement as much as possible for the problem behavior.

Differential Reinforcement of Alternative Behavior (DRA)

3. A practitioner applying DRA reinforces occurrences of a desirable alternative to the problem behavior and withholds reinforcement for the problem behavior.

4. DRA can be conceptualized as a concurrent schedule of reinforcement in which two response classes receive reinforcement at different rates: (a) the alternative behavior on a dense schedule of high-quality reinforcement and (b) the problem behavior on extinction or a very lean schedule of reinforcement.

5. The learner allocates more responding to the alternative behavior and less responding to the problem behavior, as described by the matching law.

6. When DRA entails reinforcing a behavior that cannot occur simultaneously with the problem behavior, the procedure is sometimes called differential reinforcement of incompatible behavior (DRI).

7. When using DRA, practitioners should:
   - Select alternative behaviors that are present in the learner's repertoire, that require equal or less effort than the problem behavior, that are being emitted prior to intervention with sufficient frequency to provide opportunities for reinforcement, and that are likely to produce reinforcement when the intervention ends.
   - Select potent reinforcers that can be delivered when the alternative behavior occurs and withheld following instances of the problem behavior. The consequences that maintain the problem behavior can function as a reinforcer for the alternative behavior.
   - Reinforce the alternative behavior on a continuous reinforcement schedule initially and then gradually thin the schedule of reinforcement.
   - Be alert to increased occurrences of the problem behavior when treatment shifts from a clinical to a natural setting, when new therapists and/or caregivers begin working with the client, and when the reinforcement schedule for the alternative behavior is thinned to levels that can be maintained in the client's everyday environment.
   - Combine DRA with other reductive procedures to produce a more potent intervention.

Differential Reinforcement of Other Behavior (DRO)

8. A practitioner applying DRO delivers reinforcers contingent on the problem behavior not occurring throughout intervals of time (interval DRO) or at specific moments of time (momentary DRO).

9. With interval DRO, the omission of the problem behavior throughout a predetermined time interval produces reinforcement, and any instance of the problem behavior during the interval resets the interval and postpones reinforcement.

10. With momentary DRO, absence of the problem behavior at predetermined moments in time produces reinforcement; instances of the target behavior at other points in time do not alter when the next opportunity for reinforcement becomes available.

11. Availability of reinforcement with interval and momentary DRO procedures can occur on fixed- or variable-time schedules.

12. When using DRO, practitioners should:
    - Establish an initial DRO time interval that ensures the learner's current level of behavior will produce frequent reinforcement when the DRO contingency is applied.
    - Be careful not to reinforce inadvertently other inappropriate behaviors.
    - Deliver reinforcement at the intervals or moments in time specified by the DRO schedule contingent on the absence of the problem behavior *and* the absence of any other significant inappropriate behaviors.
    - Increase the DRO interval gradually based on decreases in the problem behavior.
    - Extend DRO in other settings and times of day after the problem behavior is substantially reduced in the treatment setting.
    - Combine DRO with other reductive procedures.

Differential Reinforcement of Low Rates of Responding (DRL)

13. A practitioner applying DRL delivers reinforcement (a) following each occurrence of the target behavior that is separated from the previous response by a minimum amount of time or (b) contingent on the number of responses within a period not exceeding a predetermined criterion.

14. DRL is used to decrease the rate of a behavior that occurs too frequently, but not to eliminate the behavior entirely.

15. Reinforcement on a full-session DRL schedule is delivered when responding during an entire instructional or treatment session is equal to or below a criterion limit.

16. On an interval DRL schedule of reinforcement, the practitioner divides the total session into a series of equal intervals of time and provides reinforcement at the end of each interval in which the number of responses was equal to or below a criterion limit.

17. Reinforcement on a spaced-responding DRL schedule follows each occurrence of the target behavior that is separated from the previous response by a minimum interresponse time (IRT).

18. When using DRL, practitioners should:
    - Not use DRL if a problem behavior needs to be reduced quickly.
    - Not use DRL with self-injurious or other violent behaviors.
    - Select the most appropriate DRL schedules: full-session or interval DRL when it is acceptable that the rate of the problem behavior reaches zero or as an initial step toward the goal of eliminating the behavior; spaced-responding DRL for reducing the rate of a behavior to be maintained in the learner's repertoire.
    - Use baseline data to guide the selection of the initial response or IRT limits.
    - Thin the DRL schedule gradually to achieve the desired final rate of responding.
    - Provide feedback to help the learner monitor the rate of responding.

KEY TERMS

differential reinforcement of alternative behavior (DRA)
differential reinforcement of incompatible behavior (DRI)
differential reinforcement of low rates (DRL)
differential reinforcement of other behavior (DRO)
fixed-interval DRO (FI-DRO)
fixed-momentary DRO (FM-DRO)
full-session DRL
interval DRL
spaced-responding DRL
variable-interval DRO (VI-DRO)
variable-momentary DRO (VM-DRO)

MULTIPLE-CHOICE QUESTIONS

1. Joni works in a group home, with a group of men with disabilities. She has a good relationship with each of the men, but one of the residents continues to make sexually-inappropriate comments to her. Joni typically replies, "Frank, stop calling me that! You know I don't like it and it's not appropriate," as she turns red in the face. An informal functional behavior assessment indicates this behavior is maintained by Joni's attention. The group home staff decides to implement a DRA intervention. Which of the following would be an appropriate example of DRA in this case?
   - **a.** Providing enthusiastic attention whenever Frank says anything appropriate to Joni, while ignoring him if he makes inappropriate remarks
   - **b.** Providing enthusiastic attention whenever Frank makes an appropriate greeting to Joni (such as "You look nice today, Joni."), while ignoring Frank if he uses inappropriate greetings
   - **c.** Providing increased enthusiastic attention to Frank throughout the day
   - **d.** Ignoring all of Frank's inappropriate verbalizations

   Hint: (See "Differential Reinforcement Defined")

2. Differential reinforcement always includes which two principles of behavior?
   - **a.** Reinforcement and punishment
   - **b.** Reinforcement and stimulus control
   - **c.** Reinforcement and extinction
   - **d.** Positive and negative reinforcement

   Hint: (See "Differential Reinforcement Defined")

3. In DRA, a practitioner:
   - **a.** Reinforces a desirable alternative to the problem behavior
   - **b.** Places the problem behavior on extinction
   - **c.** May choose to reinforce a behavior that is incompatible with the problem behavior
   - **d.** All of these are correct

   Hint: (See "Differential Reinforcement of Alternative Behavior")

4. The behavior selected as an alternative behavior in a DRA:
   - **a.** Should already be in the learner's repertoire
   - **b.** Should require more effort than the problem behavior
   - **c.** Should be emitted at a very low rate prior to intervention
   - **d.** Should require extensive training for practitioners to learn how to reinforce it

   Hint: (See "Differential Reinforcement of Alternative Behavior")

5. DRO interventions:
   - **a.** Provide reinforcement for a specific alternative behavior
   - **b.** Provide reinforcement for slowly increasing occurrences of a behavior
   - **c.** Provide reinforcement for incompatible behaviors
   - **d.** Provide reinforcement for not responding

   Hint: (See "Differential Reinforcement of Other Behavior")

6. A limitation of DRO is:
   - **a.** Other inappropriate behaviors may be reinforced accidentally
   - **b.** One may accidentally punish all responding in the individual.
   - **c.** It often produces an extinction burst effect.
   - **d.** The effects have been shown to be slow and gradual.

   Hint: (See "Differential Reinforcement of Other Behavior")

7. A DRL intervention is useful when the practitioner wants to:
   - **a.** Get rid of a target behavior all together
   - **b.** Teach a new response to replace the problem behavior
   - **c.** Decrease the overall rate of a problem behavior, but not get rid of it all together
   - **d.** Increase the overall rate of a target behavior

   Hint: (See "Differential Reinforcement of Low Rates of Responding")

8. If the mean interresponse time (IRT) of a target problem behavior is 12 responses per hour, and you are implementing a DRL procedure, what should the initial DRL criterion be?
   - **a.** More than 12 responses per hour
   - **b.** No more than 11 responses per hour
   - **c.** No more than 1 response per hour
   - **d.** No more than 4 responses per hour

   Hint: (See "Differential Reinforcement of Low Rates of Responding")

ESSAY-TYPE QUESTIONS

1. The introduction to this chapter states that the procedures described in this chapter are reinforcement-based procedures for decreasing problem behavior. But, doesn't any reinforcement-based procedure, by definition, increase a target behavior and not decrease a target behavior? Explain how a differential reinforcement procedure can be used to decrease a target behavior.

   Hint: (See "Differential Reinforcement Defined")

2. When implementing a DRA intervention, it is useful to select reinforcers that are powerful and can be delivered consistently. Explain the importance of selecting reinforcers for such interventions, and explain the best method for selecting reinforcers for such interventions.

   Hint: (See "Differential Reinforcement of of Alternative Behavior")

3. Explain the difference between a momentary and interval DRO and when you might use each.

   Hint: (See "Differential Reinforcement of Other Behavior")

4. Explain the difference between full-session, interval, and spaced-responding DRL and when you might use each.

   Hint: (See "Differential Reinforcement of Low Rates of Responding")

NOTES

1. Differential reinforcement is also used to increase desired behavior, as in the differential reinforcement of high rates (DRH) (see Chapter 13) and shaping (see Chapter 22). Various forms of differential reinforcement contingencies are also used as experimental control procedures (see Chapter 8 and Thompson & Iwata, 2005).
2. DRA is a primary component of *functional communication training (FCT)*, an intervention in which the alternative behavior serves the same communicative function as the problem behavior (e.g., saying "Break, please" results in the same reinforcement as aggression or tantrums did previously). FCT is described in detail in Chapter 26.
3. Because instances of target problem behavior in DRO and DRL (to be discussed later in this chapter) often postpone reinforcement, some behavior analysts conceptualize DRO and DRL as negative punishment procedures rather than reinforcement techniques. For example, Van Houten and Rolider (1990) consider DRO and DRL variations of a procedure they called *contingent reinforcement postponement*.
4. DRO can also be applied using data from permanent product measurement. For example, Alberto and Troutman (2013) described a procedure in which a teacher provided reinforcement each time a student submitted a paper that contained no doodles. If the papers were basically of the same length or type, the procedure would be a variation of the fixed-interval DRO.
5. Hammond and colleagues (2011) reported several examples of different signals used with FM-DRO: *turning away, prompts,* and *praise* (Harris & Wolchik, 1979), *gesture* (Repp, Barton, & Brulle, 1983), *tone* (Conyers, Miltenberger, Romaniuk, Kopp, & Himle, 2003), and *music* (Sisson, Van Hasselt, Hersen, & Aurand, 1988).
6. Because reinforcement in a spaced-responding DRL immediately follows an instance of the target behavior, it is the applied variation of DRL that most closely resembles the DRL schedules of reinforcement described by Ferster and Skinner (1957).

CHAPTER 26

Antecedent Interventions

LEARNING OBJECTIVES

- Define and discuss applications of antecedent interventions.
- Define and discuss applications of noncontingent reinforcement.
- Define and discuss applications of high-probability request sequences.
- Define and discuss applications of functional communication training.

The owner of a private elementary school decided to dedicate her school to the applications of direct instruction, precision teaching, and behavior analysis. The school served students with mild to moderate intellectual disabilities, learning disabilities, and problem behaviors. To assist with the school's change of focus, the owner contacted a professor at a local university with expertise in applied behavior analysis to conduct weekly after-school in-service trainings for her teachers. The school's owner wanted the in-services to concentrate on (a) principles of behavior, (b) measurement and analysis, and (c) strategies and tactics for application.

The in-service progressed well. The teachers developed a school-wide token economy that included a store where students could exchange their tokens based on their performance. However, the teachers reported a growing problem during recess—four boys were hitting and shoving other students.

In subsequent discussions with the teachers, several intervention approaches were presented, including response cost or time-out, to address the inappropriate playground behavior. One teacher made this suggestion.

Let's not use punishment with these boys. I suggest we develop a playground safety patrol, with the four boys serving as the patrol officers, and expand the use of token economy during recess. We could meet with the boys, tell them about the new safety patrol, and that they have been selected to serve as the first patrol officers. Then, we give them their safety patrol badges to wear. We will need to teach the new officers (a) how to identify good playground behavior—we can describe examples of good playground behavior, and role-play this playground behavior; (b) how to praise good playground behavior; and, finally, (c) how to present tokens for good playground behavior of other children during recess.

The group embraced this option and implemented it. Within 2 weeks of implementation, inappropriate playground behavior decreased, and the teachers gradually faded out the presence of the playground safety patrol.

Anecdotally, the teachers reported that whereas the four boys did not have many positive playground interactions with other students before the antecedent intervention, they appeared to be making friends on the playground subsequent to it.

In the early days of applied behavior analysis, researchers and practitioners emphasized the three-term contingency: how consequences affect behavior and how differential consequences produce stimulus discrimination and stimulus control. Seldom did applied behavior analysts attempt to change behavior by manipulating antecedent events. This situation changed dramatically following three groundbreaking publications in the 1980s: Jack Michael's (1982) conceptual analysis of establishing operations; Brian Iwata and colleagues' demonstration of a method to assess functional relations between self-injury and specific antecedent events (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982); and Ted Carr and Mark Durand's (1985) experiment showing that teaching children to emit an appropriate communicative response during conditions that previously occasioned disruptive behavior reduced misbehavior to near zero. Building upon the concepts and techniques introduced in these discipline-altering articles, behavior analysts have developed a wide array of antecedent interventions that diminish problem behaviors and promote adaptive behaviors in social, academic, leisure, and work environments.

DEFINING AND CLASSIFYING ANTECEDENT INTERVENTIONS

An **antecedent intervention** is implemented *prior to* and *independent of* the target behavior's occurrence. Some antecedent interventions manipulate motivational variables, some make the target behavior more or less effortful, some entail training an alternative behavior, some include differential consequences for responding, and some change the environment such that the opportunities to engage in a problem behavior are restricted (or,

conversely, opportunities made more prevalent for a desired behavior). Table 26.1 shows a sampling of antecedent interventions practitioners might use to address common behavior problems and promote positive behavior.

Although identifying such varied interventions with a single term is economical, recognizing the basic principles at the heart of each intervention is the first step toward effective application. Antecedent interventions based on stimulus control (discriminative stimuli [S^Ds]) evoke behavior because they have been correlated with increased availability of reinforcement. The evocative function of motivating operations (MOs), however, is independent of the differential availability of reinforcement. For example, an establishing operation will increase the current rate of certain types of behavior even when an effective reinforcer is not available. Failure to distinguish the different functions of antecedent events leads to conceptual confusion and ineffective, or possibly even unethical, service to clients.

In addition to improved conceptual clarity and consistency, understanding the different reasons for the evocative functions of S^Ds and MOs has important applied implications. Antecedent treatments involving stimulus control must include manipulating the differential availability of reinforcement in the presence and absence of the S^D. Behavior change strategies based on motivating operations must change antecedent events. Understanding these differences may improve the development of more effective and efficient behavior change efforts involving antecedent events.

Classifying Antecedent Interventions

Antecedent interventions can be classified in several ways. For example, Smith (2011) identified two broad categories of antecedent interventions for problem behavior. "*Function-based antecedent interventions* directly manipulate at least one component of a particular operant contingency that has been identified to maintain the problem behavior" (p. 297). With this definition, antecedent interventions based on stimulus control and those that involve motivating operations would be considered function-based. Smith classified antecedent-based interventions "whose effects do not depend on specific identification of the variables that set the occasion for and maintain the problem behavior" as *default interventions* (p. 297).

Antecedent interventions can also be classified by whether or not differential consequences are required for behavior change. The effects of a *contingency-dependent* antecedent intervention depend on differential consequences for the target behavior (or alternative behavior) in the presence or absence of the antecedent. Antecedent interventions based on stimulus control are contingency dependent. For example, in the presence of 2 + 2 = ?, a student responds 4, not because of the stimulus 2 + 2 = ?, but because of a past reinforcement history for saying 4, including perhaps a history of non-reinforcement for responses other than 4. Chapter 17 addresses contingency-dependent antecedent interventions based on stimulus control.

The effects of a *contingency-independent* antecedent intervention do not depend on differential consequences for the target behavior. Smith's (2011) default antecedent interventions are contingency independent, as are antecedent interventions based on motivating operations. The evocative and abative effects of establishing and abolishing operations occur independent of differential consequences (see Chapter 16). Table 26.2 provides examples of abolishing operations to decrease the effectiveness of the reinforcer maintaining various problem behaviors and a corresponding reduction of those behaviors. Chapter 16 discusses motivating operations to effect behavior change.

Smith and Iwata (1997) remind us that the effects of MOs are temporary and may not produce permanent improvements in behavior. However, practitioners can supplement MO manipulations with extinction and differential reinforcement to change behavior on a more permanent basis (Phillips, Iannaccone, Rooker, & Hagopian, 2017).

The next sections of this chapter describe three contingency-independent antecedent interventions with established scientific evidence of their effectiveness: noncontingent reinforcement, high-probability instructional sequence, and functional communication training. The chapter closes with brief descriptions of three default antecedent interventions: restraint, enriched environment, and antecedent exercise.

TABLE 26.1 Examples of Challenging Behavior and Antecedent Interventions That Might Address Them

Challenging Behavior Situation	Antecedent Intervention
When directed to do homework or take a bath, the child is noncompliant.	Parents provide a choice: "Do you want to do your homework first or take a bath first?"
When requested to complete a math worksheet with 25 problems, student disrupts class.	Teacher breaks 25 problems into 5 worksheets of 5 problems each.
After entering the classroom and waiting for the first lesson to begin, several students misbehave.	Teacher pins an index card for each student on the bulletin board. Each card contains a personalized question. Upon entering the classroom, students remove their cards from the bulletin board, go to their desks, and write an answer on the card.
After staying on task for several minutes, a nonvocal verbal employee terminates the work activity by crying and throwing materials to the floor.	Supervisor teaches the employee to sign "Break please" when tired or frustrated, and provides a brief break when the employee does so.
While creeping and crawling around the house, an infant may encounter many potentially dangerous events.	Attach gates to stairways, plug electrical outlets, put locks on cabinet doors, and remove table lamps.

TABLE 26.2 Examples of Antecedent Interventions Using Abolishing Operations

Abolishing Operation	Example
Provide corrective prompts as an antecedent event.	Antecedent corrective academic prompts reduced destructive behavior to zero (Ebanks & Fisher, 2003).
Provide presession exposure to stimuli that function as reinforcers.	A father–son playtime preceding compliance sessions improved son's compliance with father's requests (Ducharme & Rushford, 2001).
Provide free access to leisure activities.	Manipulation of leisure items effectively competed with self-injurious behavior (SIB) maintained by automatic reinforcement (Lindberg, Iwata, Roscoe, Worsdell, & Hanley, 2003).
Reduce noise levels.	Reducing noise levels decreased stereotypic covering of ears with hands (Tang, Kennedy, Koppekin, & Caruso, 2002).
Change levels of social proximity.	Low levels of distant proximity reduced aggressive behaviors (Oliver, Oxener, Hearn, & Hall, 2001).
Offer choices.	Escape-maintained problem behaviors decreased when students had opportunities to choose among tasks (Romaniuk et al., 2002).
Increasing response effort.	Increasing response effort for pica produced reductions in pica (Piazza, Roane, Kenney, Boney, & Abt, 2002).

NONCONTINGENT REINFORCEMENT

Noncontingent reinforcement (NCR) is an antecedent intervention in which stimuli with known reinforcing properties are delivered on a fixed-time (FT) or variable-time (VT) schedule independent of the learner's behavior.[1] Noncontingent reinforcement may effectively diminish difficult-to-manage behaviors because the reinforcers that maintain the behavior are available freely and frequently, thereby functioning as an abolishing operation (AO) that reduces the motivation to engage in the behavior. NCR is an important and effective antecedent intervention for the treatment of severe problem behavior displayed by persons with intellectual and/or developmental disabilities (Phillips et al., 2017).

A functional analysis of the reinforcing contingency maintaining the problem behavior determines whether NCR is implemented with positive reinforcement, negative reinforcement (escape), or automatic reinforcement.

NCR with Positive Reinforcement

Kahng, Iwata, Thompson, and Hanley (2000) provided an excellent example of applying NCR with positive reinforcement. A functional analysis showed that social-positive reinforcement maintained self-injurious behavior (SIB) or aggression in three adults with developmental disabilities. During baseline, each occurrence of SIB or aggression produced attention for two of the adults and a small bit of food for the third adult. During the initial sessions of NCR, the adults received attention or small bits of food on an initial fixed-time (FT) schedule (e.g., 5 seconds). Later, the schedule was thinned to a terminal criterion of 300 seconds. Figure 26.1 presents the baseline and NCR performances of the three participants and shows that the NCR procedure effectively decreased occurrences of SIB and aggression.

NCR Escape

Kodak, Miltenberger, and Romaniuk (2003) analyzed the effects of NCR escape on the instructional task compliance and problem behaviors of Andy and John, 4-year-old boys with autism. Andy's task was to point to cards with teacher-specified pictures, words, or letters on them. John used a marker to trace each letter of written words. Problem behaviors included resisting prompts, throwing materials, and hitting. During baseline, the therapist gave an instruction for task engagement and, contingent on problem behavior following the instruction, removed the task materials and turned away from the child for 10 seconds. During the NCR escape condition, the therapist used an initial FT 10-sec schedule for escape, meaning that the student had a break from the instructional requests every 10 seconds of the session. The initial 10-sec FT schedule was thinned each time the boy achieved a criterion for two consecutive sessions: 10 seconds to 20 seconds, to 30 seconds, to 1 minute, to 1.5 minutes, and finally to a terminal criterion of 2 minutes. The NCR escape procedure increased compliance and decreased problem behaviors.

NCR with Automatic Reinforcement

Lindberg, Iwata, Roscoe, Worsdell, and Hanley (2003) used NCR as a treatment to decrease the self-injurious behavior (SIB) of two women with profound intellectual disability. A functional analysis documented that automatic reinforcement maintained their SIB. The NCR procedure provided Julie and Laura with free access to a variety of home-based, highly preferred leisure items (e.g., beads, string) that they could manipulate throughout the day. Figure 26.2 shows that NCR object manipulation of preferred leisure items effectively diminished SIB, and the effects were maintained up to a year later. This experiment is important because it showed that NCR object manipulation could compete with automatic reinforcement to reduce the occurrence of SIB. However, treating problem behaviors maintained by automatic reinforcement often will require adding other treatments (e.g., response blocking) to accomplish clinically important effects (Phillips et al., 2017).

Using NCR Effectively

The following procedural recommendations identify three key elements for enhancing the effectiveness of NCR. (a) The

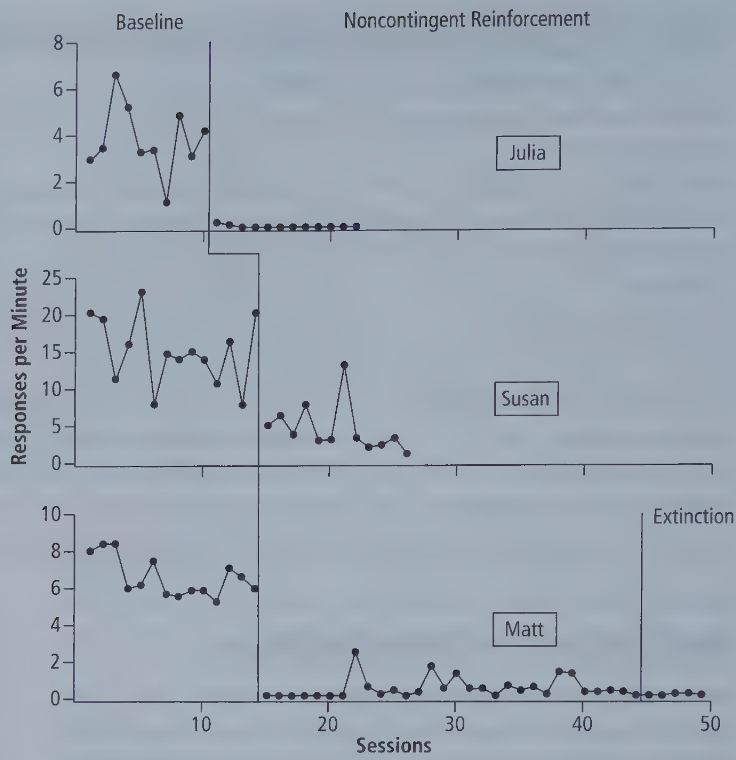


Figure 26.1 Number of SIB or aggressive responses per minute during baseline and NCR by three adults with developmental disabilities.

From "A Comparison of Procedures for Programming Noncontingent Reinforcement Schedules," by S. W. Kahng, B. A. Iwata, I. G. DeLeon, and M. D. Wallace, 2000, *Journal of Applied Behavior Analysis*, *33*, p. 426. Used by permission. Copyright 2000 by the Society for the Experimental Analysis of Behavior. Reproduced by permission.

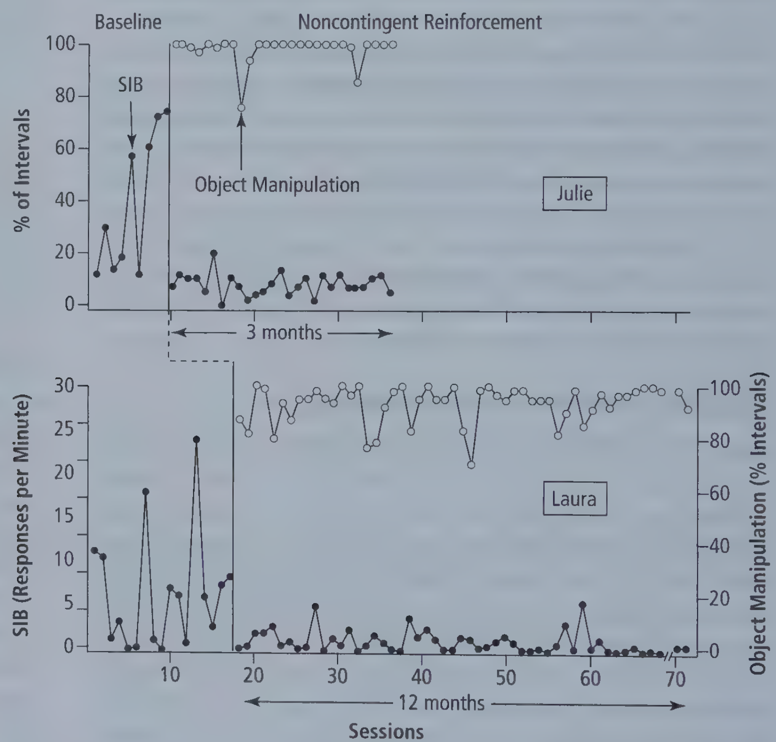


Figure 26.2 Levels of SIB and object manipulation exhibited by Julie and Laura during observations at home while NCR was implemented daily.

From "Treatment Efficacy of Noncontingent Reinforcement During Brief and Extended Application," by J. S. Lindberg, B. A. Iwata, E. M. Roscoe, A. S. Worsdell, and G. P. Hanley, 2003, *Journal of Applied Behavior Analysis*, *36*, p. 14. Copyright 2003 by the Society for the Experimental Analysis of Behavior. Reproduced by permission.

amount and quality of stimuli with known reinforcing properties influence the effectiveness of NCR. (b) Most treatments include extinction with NCR interventions. (c) Reinforcer preferences can change during intervention. That is, the NCR stimuli may not continue competing with the reinforcers that maintain the problem behavior. DeLeon, Anders, Rodriguez-Catter, and Neider (2000) recommended periodically using a variety of available stimuli with the NCR intervention to reduce problems of changing preferences.

Functional Behavior Assessment

The effectiveness of using NCR depends on the correct identification of the positive, negative, or automatic reinforcers maintaining the problem behavior. Advances in conducting

functional behavior assessments have greatly improved the effectiveness of NCR by accurately pinpointing the contingencies maintaining the behavior.[2]

Emphasizing NCR

Applied behavior analysts can enhance the effectiveness of an NCR intervention by increasing the number of known reinforcing stimuli over the non-NCR condition. For example, Ringdahl, Vollmer, Borrero, and Connell (2001) found that NCR was ineffective when the baseline condition and the NCR condition contained a similar number of reinforcers. NCR was effective, however, when the NCR schedule was denser than the baseline schedule. Applied behavior analysts can use the rates of reinforcement during baseline to establish an initial NCR schedule to ensure a discrepancy between baseline and NCR conditions.

Ringdahl and colleagues (2001) suggested three procedures for emphasizing reinforcement during the NCR intervention: (a) increase the delivery of stimuli with known reinforcing properties, (b) use a different schedule of reinforcement at treatment onset (e.g., continuous delivery or availability of stimuli with known reinforcing properties), and (c) combine differential reinforcement of other behavior (DRO) with the NCR treatment package. DRO will decrease the adventitious reinforcement of the problem behavior from the time-based NCR schedule.

Time-based NCR Schedules

Most applications of NCR use a fixed-time schedule (FT) for the delivery of stimuli with known reinforcing properties. A **fixed-time schedule** (FT) means that the interval of time for the presentation of these stimuli remains the same from delivery to delivery. When applied behavior analysts program the NCR time interval to vary across deliveries, it is called a **variable-time schedule** (VT). For example, an NCR VT schedule of 10 seconds means that, on the average, stimuli with known reinforcing properties are presented every 10 seconds. This VT schedule could use time intervals such as 5, 8, 10, 12, or 15 seconds, arranged to occur in random sequence to produce effective results (Carr, Kellum, & Chong, 2001).

Setting the initial NCR time schedule is an important aspect of the NCR procedure. The initial schedule can have an impact on the effectiveness of the intervention (Kahng, Iwata, DeLeon, & Wallace, 2000). Researchers consistently recommend an initial dense FT or VT schedule (e.g., Van Camp, Lerman, Kelley, Contrucci, & Vorndran, 2000). The therapist can set a dense time value (e.g., 4 seconds) arbitrarily. Usually, however, it is more effective to set the initial time value based on the number of occurrences of the problem behavior, which will ensure frequent contact with the NCR stimuli.

The following procedure can be used to determine an initial NCR schedule: Divide the total duration of all baseline sessions by the total number of occurrences of the problem behavior recorded during baseline, and set the initial interval at or slightly below the quotient. For example, if the participant emitted 300 aggressive acts during 5 days of the baseline, and each baseline session was 10 minutes in duration (i.e., 600 seconds), then 3000 seconds divided by 300 responses produces a quotient of 10 seconds. Accordingly, these baseline data suggest an initial FT interval of 7 to 10 seconds. Analysts then thin the schedule by adding small time increments to the NCR interval. However, thinning a time-based schedule is best begun only after the initial NCR interval has produced a reduction in the problem behavior.

Applied behavior analysts have used three procedures to thin NCR schedules: (a) constant time increase, (b) proportional time increase, and (c) session-to-session time increase or decrease (Hanley, Iwata, & Thompson, 2001; Van Camp et al., 2000). A therapist, for example, can increase the FT or VT schedule intervals by using a constant duration of time, and decrease the amount of time that the learner has access to the NCR stimuli by a constant time factor. So, a therapist might increase the schedule interval by 7 seconds at each opportunity, and each time decrease access to the stimuli by 3 seconds.

Likewise, a therapist can increase the FT or VT schedule intervals proportionately, meaning that each interval is increased by the same percentage of time. With this procedure, each time interval might be increased by 5%. If the initial time interval was 600 seconds, then the next interval would be 630 seconds ($5\% \times 600 = 30$ seconds; $600 + 30 = 630$ seconds) and so on.

Finally, a therapist can use the learner's performance to change the schedule interval on a session-to-session basis. For example, at the end of a session, the therapist establishes a new NCR time interval for the next session by dividing the number of problem behaviors that occurred in that session by the duration of the session and using that quotient as the next session's FT interval.

A therapist should adjust the interval if the problem behavior worsens during schedule thinning. The duration of the NCR interval can be readjusted after control of the problem behavior has been re-established. The readjustment, however, should proceed more gradually.

Setting Terminal Criteria

Applied behavior analysts usually select an arbitrary terminal criterion for NCR schedule thinning. Kahng and colleagues (2000) concede that research has not established a de facto terminal criterion for NCR schedule thinning. While a variety of time intervals have been used—3 minutes, 5 minutes, 10 minutes—it seems that a 5-min FT schedule has found the most favor in applied settings. The 5-min FT schedule seems to be both practical and effective.

Box 26.1 describes guidelines for using NCR in the classroom suggested by Coy and Kostewicz (2018).

Considerations for Using NCR

NCR makes an effective intervention. It has advantages in addition to effectiveness, and some disadvantages. Table 26.3 lists the advantages and disadvantages of using NCR.

BOX 26.1

Implementing NCR in the Classroom

1. *Establish context.* The first step involves using data collection and observation to identify and place a well-defined problem behavior in context. Definitions should avoid large categories of behavior and the absence of behavior, such as "off task," "disruptive," or "not listening." Behavioral definitions should include examples and nonexamples; for instance, calling out without permission includes all instances when the student yells or says something without being called on but does not include instances of choral responding. After defining the behavior, teachers then determine what commonly occurs before and after the target behavior. This contextual information provides better insight into why the behavior occurs. In addition to anecdotal or observational information, more formalized methods can provide a well-rounded picture of the target behavior. After determining the context surrounding the target behavior, teachers collect baseline data on the target behavior, collecting data at times when the behavior is most likely to occur. Typical types of data collection include (a) frequency (how many times the behavior occurs within a certain amount of time), (b) duration per occurrence (how long the behavior occurs), (c) latency (how long it takes for the behavior to initiate following a direction or request), or (d) interresponse time (IRT; time between the end of one instance of the behavior and the beginning of another).
2. *Identify potential reinforcers.* NCR relies on careful and purposeful use of positive reinforcers (e.g., items or events the student enjoys or will work for). The reinforcer should match the common consequence for the target behavior. Using a functional reinforcer (such as a break from task demands for escape-maintained behaviors or praise for attention-seeking behaviors) gives NCR the greatest chance of success.
3. *Deliver reinforcers.* During NCR, the functional reinforcer should be provided on a continuous basis and not in response to the occurrence of the target behavior. Teachers enrich the environment by providing reinforcement more often than the student currently receives for the problem behavior.
4. *Finalize the plan.* Any formal behavioral intervention must be well organized prior to implementation. Written behavior plans include background information, current levels of performance, implementation guidance, and monitoring or evaluation sections. NCR interventions maintain the same requirements.
5. *Implement and evaluate.* After reviewing the student's baseline behavior data and completing the NCR Planning Worksheet, it is time to put the plan into action. Practitioners implementing NCR must follow the schedule closely and carefully. Electronic technologies (e.g., cell phone applications) can assist with the implementation of NCR. Teachers should be aware of a few considerations when implementing NCR. The student may engage in the problem behavior just before the end of an interval. Instead of immediately providing the reinforcer, withhold the presentation of the reinforcer for approximately 10 seconds or until the student no longer displays the target problem behavior. Once the reinforcer is provided, the next NCR interval starts. Practitioners should also be prepared for brief escalations in the problem behavior during NCR. The teacher monitors yet ignores these escalations and provides the reinforcer according to the current schedule of reinforcement.

Adapted from "Noncontingent Reinforcement: Enriching the Classroom Environment to Reduce Problem Behaviors," by J. N. Coy and D. E. Kostewicz, 2018, *Teaching Exceptional Children*, *50*, 301–309. Copyright 2018 by Council for Exceptional Children. Used by permission.

TABLE 26.3 Possible Advantages and Disadvantages of Noncontingent Reinforcement

Advantages

- NCR is easier to apply than positive reductive techniques that require monitoring the client's behavior for the contingent delivery of the reinforcer (Phillips et al., 2017).
- NCR helps create a positive learning environment, which is always desirable during treatment.
- A package treatment that includes NCR with extinction may reduce extinction-induced response bursts (Van Camp et al., 2000).
- Chance pairings of appropriate behavior and NCR delivery of stimuli with known reinforcing properties could strengthen and maintain those desirable behaviors (Roscoe, Iwata, & Goh, 1998).

Disadvantages

- Free access to NCR stimuli may reduce motivation to engage in adaptive behavior.
- Chance pairings of problem behavior and NCR delivery of stimuli with known reinforcing properties could strengthen the problem behavior (Van Camp et al., 2000).
- NCR escape can disrupt the instructional process.

HIGH-PROBABILITY INSTRUCTIONAL SEQUENCE

Student compliance provides opportunities for the development of many important behaviors. Noncompliance, however, is a prevalent problem with persons who have developmental disabilities and behavior disorders. The high-probability instructional sequence is a nonaversive procedure for improving compliance and diminishing escape-maintained problem behaviors. The high-probability instructional sequence may decrease excessively slow responding to instructions and thereby reduce overall time for completing tasks (Mace et al., 1988).

When using a **high-probability instructional sequence (high-*p*)**, the practitioner presents two to five easy-to-follow instructions for which the participant has a history of compliance. When the learner complies with several high-*p* instructions in sequence, the practitioner immediately gives the target instruction (i.e., low-*p*). Engelmann and Colvin (1983) provided one of the first formal descriptions of the high-*p* instructional sequence in their compliance training procedure for managing severe behavior problems. They used the term *hard task* in the following dressing anecdote.

Typical Instructional Sequence

Teacher: "Please put on your shirt." (low-*p* request)

Student: Throws a tantrum to avoid or escape the hard task request.

Improved Instructional Sequence with Embedded High-*p* Instructional Sequence

Teacher: "Give me five!" (high-*p* request)

Student: Slaps teacher's outstretched hand appropriately.

Teacher: "All right, nice job! Now, take this ball and put it in your pocket." (high-*p* request)

Student: Puts the ball in pocket.

Teacher: "Great! That is right! Now, please put on the shirt." (low-*p* request)

Student: Puts on the shirt without the tantrum. (p. 13, text in parentheses added)

The reductive effects of the high-*p* instructional sequence on noncompliance resemble the abative effects of an AO that reduces the current potency of reinforcement for noncompliance with the low-probability requests (i.e., reducing the value of escape from requests). In effect, "[d]uring the high-*p* sequence, compliance with several high-*p* instructions, and receipt of reinforcement contingent upon this compliance, may therefore increase compliance with the subsequent low-*p* instruction" (Lipschultz & Wilder, 2017, p. 424). Some behavior analysts use **behavioral momentum** to describe a behavior's resistance to change following an alteration in reinforcement conditions, and the term is sometimes used to describe the effects produced by the high-*p* instructional sequence (Mace et al., 1988; Nevin, Mandell, & Atak, 1983).[3]

Research has demonstrated the high-*p* instructional sequence to be an effective treatment for a wide variety of behavior problems exhibited by participants ranging from preschoolers to adults, with and without disabilities. Behavioral improvements produced by the high-*p* instructional sequence have included compliance by preschoolers (Normand, Kestner, & Jessel, 2010; Wilder, Majdalany, Sturkie, & Smeltz, 2015), compliance by school-age students with teachers' academic instructions (Axelrod & Zank, 2012; Lee, Belfiore, Scheeler, Hua, & Smith, 2004), food acceptance by children with pediatric feeding disorders (Dawson et al., 2003; Ewry & Fryling, 2016; Patel et al., 2006), increased social interactions (Jung, Sainato, & Davis, 2008), and compliance with medical examination requests by children with autism (Riviere, Becquet, Peltret, Facon, & Darcheville, 2011).

Killu, Sainato, Davis, Ospelt, and Paul (1998) evaluated the effects of the high-*p* instructional sequence on compliant responding to low-*p* instructions and occurrences of problem behaviors by three preschool children with developmental delays. The researchers used a compliance criterion of 80% or higher for selecting high-*p* requests for two children and a 60% criterion for the third child. Compliance of less than 40% was used to select the low-*p* instructional.

The request sequence began with the experimenter or a trainer presenting three to five high-*p* requests. When a child complied with at least three consecutive high-*p* requests, a low-*p* request was immediately presented. Praise was provided immediately following each compliant response. Figure 26.3 shows the children's performances before, during, and after the high-*p* sequence. The sequence delivered by two different trainers increased compliant responding to the low-*p* requests of the three children. Compliant responding maintained across time and settings.

Using the High-*p* Instructional Sequence Effectively

Select High-p Instructions from Current Repertoire

The tasks selected for high-*p* instructions should be in the learner's current repertoire, occur regularly, and have a very short duration. Ardoin, Martens, and Wolfe (1999) selected high-*p* instructions by (a) creating a list of requests that corresponded to student compliance, (b) presenting each request on the list for five separate sessions, and (c) selecting as high-*p* requests only those tasks that the student complied with 100% of the time. A criterion of 80% compliance or greater for high-*p* instructions is commonly used (e.g., Belfiore, Basile, & Lee, 2008; Mace et al., 1988).

Axelrod and Zank (2012) empirically identified high-*p* and low-*p* instructions for Charles and Thomas, two fifth-grade students who emitted high rates of noncompliance and received special education services in a general education classroom. The students' teacher created a list of 40 typical classroom instructions and then presented each student each instruction 10 times randomly during reading and seatwork sessions over a period of 10 days. Instructions with 80% or greater compliance were categorized as high-*p*, and those with 40% compliance or less were categorized as low-*p*. Commands complied with between 40% and 80% of the

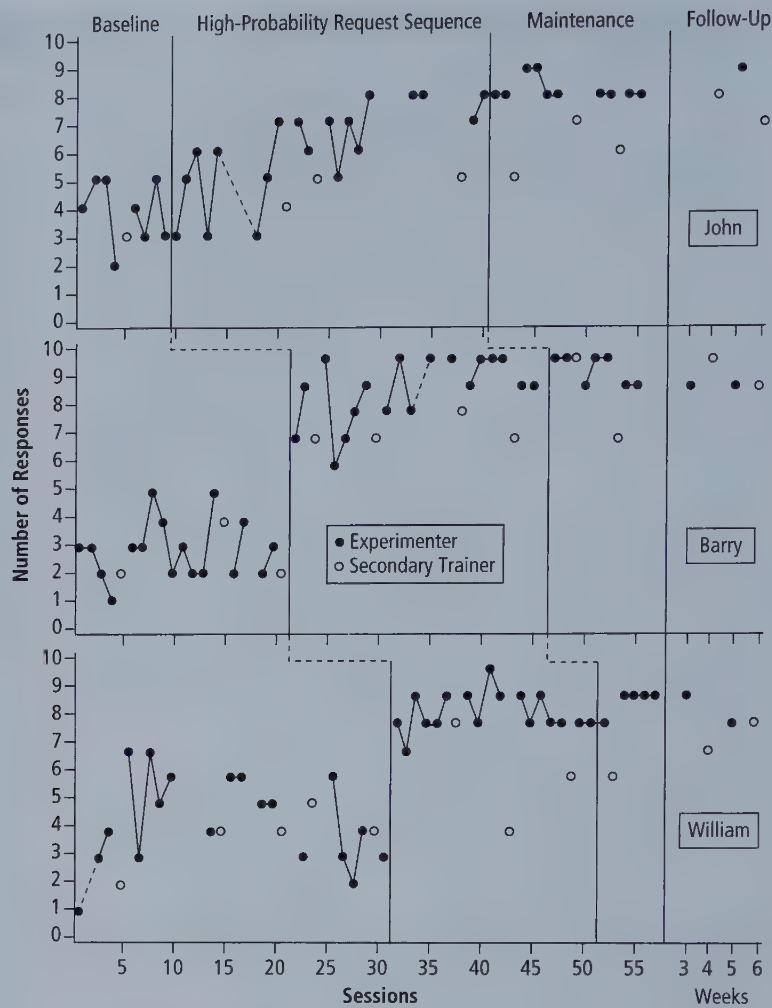


Figure 26.3 Number of compliant responses to low-probability requests delivered by the investigator and second trainer across sessions and conditions. Participants were given 10 low-*p* requests each session. Dashed lines indicate student absences.

From "Effects of High-probability Request Sequences on Preschoolers' Compliance and Disruptive Behavior," by K. Killu, D. M. Sainato, C. A. Davis, H. Ospelt, and J. N. Paul, 1998, *Journal of Behavioral Education, 8*, p. 358. Used by permission.

time were not used in the study. Table 26.4 shows Thomas' percentage compliance with teacher commands.

Mace (1996) reported that the effectiveness of the high-*p* sequence correlates with the number of increases in high-*p* requests. A high-*p* sequence with five requests may be more effective than a sequence with two requests; but the increase in effectiveness may have a trade-off in efficiency. For example, if the same, or nearly the same, effectiveness can be obtained with two or three high-*p* instructions as can be obtained with five or six, a teacher might select the shorter sequence because it is more efficient. When participants consistently comply with the low-*p* requests, the trainer should gradually reduce the number of high-*p* requests.

Present Requests Rapidly

The high-*p* requests should be presented in rapid succession, with short inter-request intervals. The first low-*p* request should immediately follow the reinforcer for high-*p* compliance (Davis & Reichle, 1996).

Acknowledge Compliance

The learner's compliance should be acknowledged immediately. Notice how the teacher in the previous dressing example acknowledged and praised the student's compliance ("All right, nice job!") before presenting the next request.

Use Potent Reinforcers

Individuals may emit aggression and self-injurious behaviors to escape from low-*p* request demands. Mace and Belfiore (1990) cautioned that social praise may not increase compliance if motivation for escape behavior is high. Therefore, high-quality positive stimuli immediately following compliance will increase the effectiveness of the high-*p* intervention.

Table 26.5 summarizes guidelines for using the high-*p* instructional sequence.

FUNCTIONAL COMMUNICATION TRAINING

Functional communication training (FCT) establishes an appropriate communicative behavior to compete with problem behaviors evoked by a motivating operation (MO). In contrast to NCR and the high-*p* instructional sequence that alter current MOs, FCT develops communicative behaviors that are sensitive to existing MOs. The practitioner uses differential reinforcement of alternative behavior (DRA) (see Chapter 25) to teach an adaptive response that produces the same reinforcer that has maintained the problem behavior, thereby making the alternative behavior functionally equivalent to the problem behavior. The alternative communicative responses can take many forms, such as vocalizations, signs, communication boards, word or picture

TABLE 26.4 Thomas' Percentage Compliance with Teacher Commands During Identification Period

Instruction Category	Instruction	Compliance During Identification Period (%)
High-*p*	Give teacher high five	100
	Give teacher fist pound	100
	Put hands on desk	90
	Pick up writing utensil (e.g., pencil, pen, color pencil)	90
	Put writing utensil in desk	80
	Put down writing utensil	80
	Put hands on lap	80
	Stand up	80
	Slide chair closer to desk	80
	Walk over to group	80
	Sit down with group	80
Low-*p*	Stop talking and continue working on assignment	10
	Put inappropriate materials (e.g., toy, drawing paper, color pencil box) in desk	10
	Stop talking and raise your hand	20
	Walk back to desk after unauthorized leave	20
	Read sentence or passage from text	30
	Write name on paper	30
	Begin working on assignment	30
	Take writing utensil out of mouth	30
	Walk over to the teacher's desk	40

Based on "Increasing Classroom Compliance: Using a High Probability Command Sequence with Noncompliant Students" by M. J. Axelrod and A. J. Zank, 2012, *Journal of Behavioral Education*, *21*, p. 125.

TABLE 26.5 Guidelines for Using the High-*p* Instructional Sequence

- Empirically identify high-*p* and low-*p* instructions before using the procedure. Compliance with high-*p* instructions should be 80% or greater; with low-*p*, 40% or less.
- Do not use the high-*p* request sequence just after an occurrence of the problem behavior. The student might learn that responding to a low-*p* request with the problem behavior will produce a series of easier requests.
- Present the high-*p* request sequence at the beginning and throughout the instructional period to reduce the possibility of problem behaviors producing reinforcement.
- Present high-*p* instructions rapidly, with brief intertrial intervals (1–5 sec) and reinforce compliance with each high-*p* instruction.
- If the participant does not comply with high-*p* instructions, stimuli associated with the low-*p* instruction should be identified and eliminated, or other high-*p* instructions should be used.
- Present first low-*p* instruction immediately following compliance with three to five high-*p* instructions.
- Present high-quality reinforcers for compliance with high-*p* instructions.
- Fade the ratio of high-*p* to low-*p* instructions to 1-to-1.
- Be alert to treatment drift in which practitioners present only high-*p* instructions to avoid escape-motivated aggression and self-injury evoked by low-*p* requests.
- If the procedure is ineffective, consider adding an extra intervention component.

Sources: Adapted from Belfiore, Basile, and Lee (2008); Davis and Reichle (1996); Dawson et al. (2003); Horner, Day, Sprague, O'Brien, and Heathfield (1991); Lipschultz and Wilder (2017); Mace (1996); Mace et al. (1988); Normand, Kestner, and Jessel (2010); Penrod, Gardella, and Fernand (2012); Pitts and Dymond (2012); and Wilder, Majdalany, Sturkie, and Smeltz (2015).

cards, vocal output systems, or gestures (Brown et al., 2000; Shirley, Iwata, Kahng, Mazaleski, & Lerman, 1997).

Carr and Durand (1985) described functional communication training as a two-step process: (1) conduct a functional behavior assessment to identify the stimuli with known reinforcing properties that maintain the problem behavior and (2) use those stimuli as reinforcers to develop an alternative behavior to replace the problem behavior. FCT is an effective treatment for many problem behaviors maintained by social attention.

FCT-based interventions typically involve several behavior change tactics in addition to teaching the alternative communicative response. For example, applied behavior analysts often use a combination of response prompting, time-out, physical restraint, response blocking, redirection, and extinction to address the difficult behavior.

Durand (1999) used FCT in school and community settings to reduce the challenging behaviors of five students with severe disabilities. Durand first completed functional behavior assessments to identify the objects and activities maintaining the problem behaviors. Following the functional assessments, the students learned to use a communication device that produced digitized speech to request objects and activities identified during the functional behavior assessments. The five students reduced the occurrences of their problem behaviors in school and community settings in which they used digitized speech to communicate. Figure 26.4 presents data on the percentage of intervals of problem behaviors occurring in community settings for each student. These data are socially significant because they show the importance of teaching skills that recruit reinforcement in natural settings, thereby promoting generalization and maintenance of intervention effects.

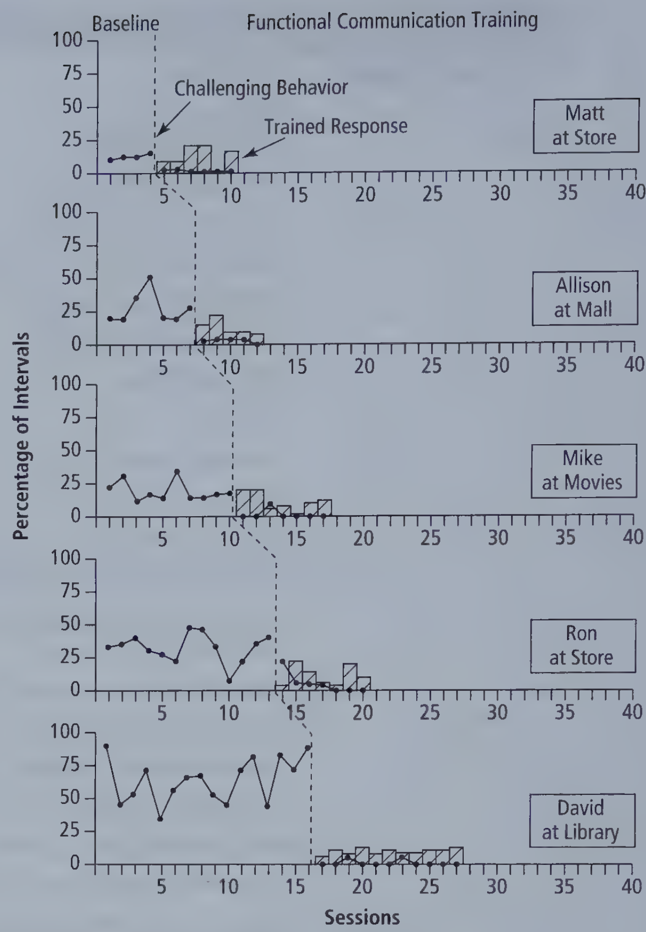


Figure 26.4 Percentage of intervals of challenging behavior by each of the five participants in baseline and FCT in community settings. The hatched bars show the percentage of intervals of unprompted communication by each student.

From "Functional Communication Training Using Assistive Devices: Recruiting Natural Communities of Reinforcement," by V. M. Durand, 1999, *Journal of Applied Behavior Analysis, 32,* p. 260. Copyright 1999 by the Society for Experimental Analysis of Behavior. Reproduced by permission.

Effective Use of FCT

A number of factors enhance the use of FCT, an appropriate and effective antecedent intervention for decreasing behavior. We discuss the principal factors in the next sections. However, we recognize Andzik, Cannella-Malone, and Sigafoos's (2016) point that few studies have analyzed practitioners' fidelity in implementing FCT. Their review sheds light on this important issue.

Dense Schedule of Reinforcement

The alternative communicative response should produce the reinforcers that maintain the problem behavior on a continuous schedule of reinforcement during the early stages of communication training.

Decreased Use of Verbal Prompts

While teaching the alternative communicative response, verbal prompts such as saying "*look*" or "*watch me*" are used often. After the communicative response is established firmly, the trainer should gradually reduce any verbal prompts and, if possible, eliminate them altogether to remove any prompt dependence associated with the intervention (Miltenberger, Fuqua, & Woods, 1998).

Behavior Reduction Procedures

The effectiveness of FCT is likely to be enhanced if it is individualized within a treatment package consisting of extinction or time-out (Shirley et al., 1997).

Schedule Thinning

Thinning the reinforcement schedule for a firmly established communicative response is an important part of the FCT treatment package. Time-based procedures, such as constant time increase, proportional time increase, and session-to-session time increase, are not appropriate for schedule thinning with the alternative communicative response. They are incompatible with the methods used to differentially reinforce the alternative communicative behavior because the FCT intervention does not alter the MO that evokes the problem behavior. The alternative communicative behavior must remain sensitive to the evocative function of the MO to compete with the problem behavior. For example, consider a child with developmental disabilities who has a history of self-stimulatory behavior when presented with a difficult task. A therapist teaches the child to ask for assistance with these tasks (i.e., the alternative communicative behavior), which reduces the self-stimulatory behavior. After firmly establishing *asking for assistance*, the therapist or caregiver should not decrease instances of providing that assistance because such a decrease might break the alternative communicative behavior–reinforcer contingency, and risk the re-emergence of self-stimulation.

Hanley, Piazza, Fisher, Contrucci, and Maglieri (2001) recommended a procedure for schedule thinning that used a dense fixed-interval schedule of reinforcement (e.g., FI 2 sec, FI 3 sec) during the initial teaching of the alternative communicative response. Once the communicative response was established, they suggested gradually thinning the FI schedule. This procedure, in contrast to time-based procedures, maintained the contingency between responding and reinforcement. They cautioned that thinning the FI schedule during FCT interventions could produce undesirable high rates of the alternative communicative response that could disrupt home or classroom settings. Hanley and colleagues further suggested the use of picture cues and external "clocks" to announce when reinforcement is available, to control the undesirable high rate of the communicative response.

For further review of research on FCT schedule thinning and recommendations for practice, see Greer, Fisher, Saini, Owen, and Jones (2016) and Hagopian, Boelter, and Jarmolowicz (2011).

Table 26.6 summarizes the advantages and disadvantages of using FCT.

DEFAULT INTERVENTIONS

Smith (2011) defined default antecedent interventions as treatments "whose effects do not depend on specific identification of the variables that set the occasion for and maintain the problem behavior" (p. 297) and described three principal types: antecedent exercise, enriched environment, and restraint. Smith and Iwata (1997) noted that antecedent exercise is used regularly for self-injury, aggression, and stereotypy; enriched environment is most frequently applied for problem behaviors maintained by automatic positive reinforcement; and restraint is an appropriate intervention for severe behavior problems in limited selected situations.

Antecedent Exercise

Antecedent exercise requires the individual to engage in some effortful form of aerobic activity (e.g., walking, jogging, dancing, calisthenics, roller skating) prior to receiving a low-*p* task to complete. Upon completion of the exercise, the individual is directed to complete the task, and performance is recorded.

Research over a few decades has reported that antecedent exercise has decreased many maladaptive behaviors, such as SIB and aggression, as well as diverse behaviors such as inappropriate vocalizations, repetitive movements, and talking-out, out-of-seat, and stereotypic behaviors. "It is important to note that in contrast to other interventions involving effortful activity, such as overcorrection (Foxx & Azrin, 1972, 1973), antecedent exercise is implemented independently of occurrences of the problem behavior" (Smith, 2011, pp. 298–299).

Celiberti, Bobo, Kelly, Harris, and Handleman (1997) used a multielement design to analyze the effects of two levels of antecedent exercise on the self-stimulatory behavior of a 5-year-old boy with autism.

> The exercise conditions were applied immediately before periods of academic programming. Maladaptive self-stimulatory behaviors were separately tracked, enabling identification of behaviors that were more susceptible to change (e.g., physical self-stimulation and "out of seat" behavior) versus those that were more resistant (e.g., visual self-stimulation). Examination of temporal effects indicated a decrease in physical self-stimulation and "out of seat" behavior, but only for the jogging condition. In addition, sharp reductions in these behaviors were observed immediately following the jogging intervention and gradually increased but did not return to baseline levels over a 40 min period (p. 139). (See Figure 26.5.)

Enriched Environment

An **enriched environment** (EE) intervention provides noncontingent access to preferred sources of reinforcement (e.g., toys, games, social and recreation activities). This noncontingent access to preferred sources of reinforcement arranges a competition between the EE intervention and the stimulation provided by the problem behavior (Horner, 1980). When effective, this competition decreases the problem behavior.

Vollmer, Marcus, and LeBlanc (1994) examined interventions for three participants following inconclusive functional analyses. These participants included Ron and Korey, 3-year-old boys with severe disabilities and self-injurious behaviors (SIB), and Rhonda, a 4-year-old-girl, also with severe disabilities, as well as chronic hand mouthing. During a stimulus preference assessment (see Chapter 11), Vollmer et al. arranged an array of teacher- or parent-suggested items

TABLE 26.6 Possible Advantages and Disadvantages of Functional Communication Training

Advantages
• Excellent chance for generalization and maintenance of the alternative communicative response because the communicative response often functions to recruit reinforcement from significant others (Fisher, Kuhn, & Thompson, 1998).
• May have high social validity. Participants report preferences for FCT over other procedures to diminish behavior (Hanley et al., 1997).
Disadvantages
• FCT treatment packages usually include extinction, which may produce undesirable effects (see Chapter 22).
• The extinction procedure is difficult to use consistently, allowing for intermittent reinforcement of problem behaviors.
• Participants may emit inappropriately high rates of the alternative communicative response (Fisher et al., 1998).
• Recruitment of reinforcement can occur at inconvenient or impossible times for the caregiver (Fisher et al., 1998).
• FCT leaves intact the environment that evoked the problem behavior, which may limit its overall effectiveness (McGill, 1999).

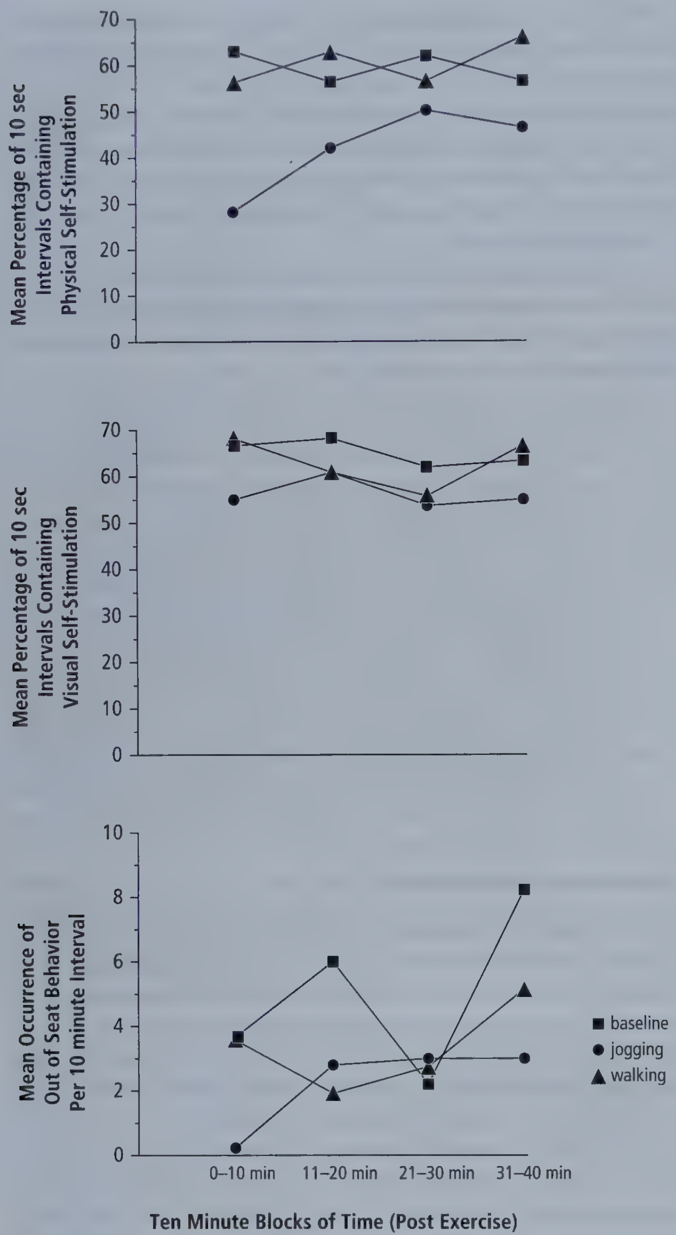


Figure 26.5 Mean percentage of intervals of physical self-stimulatory behaviors and visual self-stimulatory behaviors and mean frequency of "out of seat" behavior across four 10-min blocks of time after implementation of treatment conditions for baseline (open square), walking (open triangle), and jogging (filled circle).

From "The Differential and Temporal Effects of Antecedent Exercise on the Self-Stimulatory Behavior of a Child with Autism," by D. Celiberti, H. E. Bobo, K. S. Kelly, S. L. Harris, and J. S. Handleman, 1997, *Research in Developmental Disabilities, 18*, p. 147. Used by permission.

for each individual, including food, drinks, toys, and balls. Using standard assessment protocols, the researchers discovered the preferred and differential reinforcers for each child. For example, Rhonda preferred sound-making toys and cookies, while the other children preferred different items. The EE interventions of preferred stimuli decreased the aberrant behaviors of the three children.

Restraint

Restraint involves physically limiting, prohibiting, or securing an individual in such a manner that the target behavior cannot occur. Three forms of restraint as an antecedent intervention appear in the behavior analysis literature: personal restraint, protective equipment restraint, and self-restraint.

Personal Restraint

Personal restraint is "implemented in situations where problem behavior is highly likely to occur [and] involves caregivers physically securing and holding body parts, so that problem behavior cannot occur" (Smith, 2011, p. 300). Personal restraint as an antecedent intervention should not be confused with response blocking, which entails stopping a problem behavior that has been initiated (see Chapter 14).

Protective Equipment Restraint

Protective equipment restraints (e.g., padded helmets, arm splints, boxing gloves, padded hand mitts, safety belts, safety straps) are used to protect individuals from serious self-injury and to reduce the frequency of SIB (Smith, 2011). For example, Mazaleski, Iwata, Rogers, Vollmer, and Zarcone (1994) used oven mitts for treating the chronic hand mouthing of two women with profound intellectual disabilities. Marty was 33 years old, and Ava was 34 years old. Both lived in a public residential facility. Marty and Ava emitted chronic hand mouthing that produced mild tissue damage. In addition to hand mouthing, Ava also engaged in tongue pulling, which resulted in periodic irritation and ulcers. After collecting baseline on hand mouthing, two conditions were implemented. In a noncontingent-equipment condition, one participant wore oven mitts throughout an entire session. In a contingent-equipment condition, mitts were placed on the participant only following hand mouthing episodes. Results showed that both noncontingent and contingent mitts produced a large decrease in the rate of hand mouthing.

Self-restraint

Persons with self-injurious behavior often apply self-restraint. Wendy, a 12-year-old girl with Down syndrome, received treatment for hand-to-head hitting that caused vision loss and other facial injuries. Wendy applied self-restraint by "sitting on her hands or placing her arms between folded legs or wrapped in clothing. Wendy's caregivers reported that self-restraint was encouraged and seldom blocked, but that self-restraint interfered with adaptive skills" (Scheithauer, O'Connor, & Toby, 2015, p. 908).

The unique relationship between SIB and self-restraint raises important questions related to the controlling variables during treatment (Fisher & Iwata, 1996). Will the same reinforcement contingency maintain both behaviors? Do SIB and

self-restraint have a shared history, but different contingencies of reinforcement? Will access to SIB (or self-restraint) function as reinforcement for self-restraint (or SIB)? Research has begun to address these questions. For example, Smith, Lerman, and Iwata (1996) suggested that self-restraint provided positive reinforcement for SIB. Fisher, Grace, and Murphy (1996) proposed that self-restraint was maintained by negative reinforcement by escape from blocking SIB (e.g., pain). Derby, Fisher, and Piazza (1996) found that self-restraint and SIB were members of the same functional response class (both behaviors were maintained by social attention) and suggested noncontingent reinforcement as a possible treatment for self-restraint and SIB. Much research will be required to fully understand the variables that account for the effects of self-restraint.

SUMMARY

Defining and Classifying Antecedent Interventions

1. An antecedent intervention is implemented *prior to* and *independent of* the target behavior's occurrence.

2. Smith (2011) identified two broad categories of antecedent interventions: *Function-based antecedent interventions* manipulate at least one component of the contingency identified as maintaining the problem behavior; *default interventions* do not depend on identifying the variables that evoke and maintain the problem behavior.

3. Antecedent interventions can also be classified by whether or not differential consequences are required for behavior change. *Contingency-independent* antecedent interventions do not depend on differential consequences for the target behavior; *contingency-dependent* antecedent interventions depend on differential consequences for the target behavior (or alternative behavior) in the presence or absence of the antecedent.

Noncontingent Reinforcement

4. NCR is an antecedent intervention in which stimuli with known reinforcing properties are delivered on a fixed-time (FT) or variable-time (VT) schedule independent of the learner's behavior.

5. An NCR-enriched environment may function as an abolishing operation (AO) that reduces motivation to engage in the problem behavior.

6. A functional analysis of the reinforcing contingency maintaining the problem behavior determines whether NCR is implemented with positive reinforcement, negative reinforcement (escape), or automatic reinforcement.

High-probability Instructional Sequence

7. The basic procedure for implementing the high-probability instructional sequence (high-*p*) is as follows:
   - First, present two to five instructions for which the participant has a history of compliance and reinforce each compliant response.
   - Second, immediately after the participant has complied with several high-*p* instructions in sequence, present a target (i.e., low-*p*) instruction.
   - Third, gradually fade the ratio of high-*p* to low-*p* instructions to 1-to-1.

8. The reductive effects of the high-*p* instructional sequence on noncompliance resemble the abative effects of an AO that reduces the current potency of reinforcement for noncompliance with the low-probability requests (i.e., reducing the value of escape from requests).

9. Behavioral momentum is sometimes used to describe the effects produced by the high-*p* instructional sequence.

Functional Communication Training

10. FCT establishes an appropriate communicative behavior to compete with problem behaviors evoked by a motivating operation (MO).

11. The practitioner uses DRA to teach an adaptive response that produces the same reinforcer that has maintained the problem behavior, thereby making the alternative communicative response functionally equivalent to the problem behavior.

Default Interventions

12. Default antecedent interventions do not depend on identifying the variables that evoke and maintain the problem behavior.

13. Antecedent exercise requires the participant to engage in some form of aerobic activity prior to receiving an instruction or entering an environment in which the problem behavior has frequently occurred.

14. The noncontingent access to preferred sources of reinforcement provided by an enriched environment intervention arranges a competition between the intervention and the stimulation provided by the problem behavior.

15. Three methods of restraint—physically limiting, prohibiting, or securing an individual in such a manner that the problem behavior cannot occur—have been evaluated in the applied behavior analysis literature: personal restraint, protective equipment restraint, and self-restraint.

KEY TERMS

antecedent intervention
behavioral momentum
fixed-time schedule (FT)
functional communication training (FCT)
high-probability (high-p) request sequence
noncontingent reinforcement (NCR)
variable-time schedule (VT)

MULTIPLE-CHOICE QUESTIONS

1. Functional communication training is considered to be:
   a. Consequent intervention
   b. Punishment strategy
   c. Antecedent intervention
   d. Behavioral momentum technique

   Hint: (See "Functional Communication Training")

2. Behaviors selected for a high-p request sequence should:
   a. Be a part of the learner's current repertoire
   b. Be new behavior(s) the learner has to acquire
   c. Have a very long duration of occurrence
   d. Be those behaviors with which the learner is not compliant

   Hint: (See "High-Probability Instructional Sequence")

3. To effectively implement a high-p request sequence, it is important to:
   a. Select behaviors not in the learner's current behavior repertoire
   b. Not acknowledge compliance
   c. Use reinforcers with limited reinforcing value
   d. Present requests in a rapid sequence

   Hint: (See "High-Probability Instructional Sequence")

4. Noncontingent reinforcement involves the presentation of:
   a. Stimuli with unknown reinforcing properties that are delivered on a fixed-time or variable time schedule independent of the learner's behavior
   b. Stimuli with known reinforcing properties that are delivered on a fixed-time or variable-time schedule independent of the learner's behavior
   c. Stimuli with known reinforcing properties that are delivered on a fixed-ratio or variable-ratio schedule independent of the learner's behavior
   d. Stimuli with known reinforcing properties that are delivered on a fixed-time or variable-time schedule dependent on the learner's behavior

   Hint: (See "Noncontingent Reinforcement")

5. Noncontingent reinforcement may effectively decrease problem behavior because:
   a. Reinforcers that maintain the problem behavior are available freely and frequently.
   b. Reinforcers that maintain the problem behavior are withheld and made contingent upon the occurrence of alternative behaviors.
   c. Reinforcers that maintain the problem behavior are presented only when the problem behavior occurs.
   d. Reinforcers do not maintain the problem behavior presented regardless of the behaviors.

   Hint: (See "Noncontingent Reinforcement")

6. Functional communication training:
   a. Develops alternative behaviors that are sensitive to abolishing operations
   b. Develops alternative behaviors that create changes in motivating operations
   c. Develops alternative behaviors that are sensitive to establishing operations
   d. Develops alternative behaviors that produce punishing effects

   Hint: (See "Functional Communication Training")

7. Antecedent interventions are also known as:
   a. Antecedent procedures
   b. Antecedent control
   c. Antecedent manipulations
   d. All of these

   Hint: (See "Defining and Classifying Antecedent Interventions")

ESSAY-TYPE QUESTIONS

1. What is an antecedent intervention?
   Hint: (See "Defining and Classifying Antecedent Interventions")
2. Describe noncontingent reinforcement (NCR) and its application as an antecedent intervention strategy.
   Hint: (See "Noncontingent Reinforcement")
3. Describe high-probability (high-p) request sequence and its application as an antecedent intervention strategy.
   Hint: (See "High-Probability Instructional Sequence")
4. Describe functional communication training (FCT) and its application as an antecedent intervention strategy.
   Hint: (See "Functional Communication Training")

NOTES

1. Interpreting the NCR procedure as presenting a reinforcer is technically inconsistent with the definition of reinforcement, which requires a response-reinforcer relation (Poling & Normand, 1999). We use NCR to describe the time-based delivery of stimuli with known reinforcing properties because applied behavior analysts have continued its use, and the term NCR serves a descriptive purpose.
2. Chapter 27 provides a detailed description of functional behavior assessment.
3. To learn more about behavioral momentum theory describing response rate and resistance to change as two separate aspects of operant behavior, see Craig, Nevin, and Odum (2014); Greer, Fisher, Romani, and Saini (2016); Nevin (1992); and Shahan and Sweeney (2011).

PART 10

Functional Assessment

Changing behavior is often challenging, perplexing, and frustrating. The behavior analyst's task of determining what to do and how to do it can be particularly difficult when helping people with chronic problem behaviors that have been resistant to change. In Chapter 27 Stephanie Peterson and Nancy Neef describe an assessment process that explores the function (or purpose) a behavior serves for a person. Functional behavior assessment enables behavior analysts to make empirically based hypotheses for *why* problem behaviors occur, information that can point to the design of effective interventions. Peterson and Neef describe the basis for functional behavior assessments and their role in treating and preventing problem behaviors, and present case examples illustrating three methods for conducting functional behavior assessments.

CHAPTER 27

Functional Behavior Assessment

Stephanie M. Peterson and Nancy A. Neef

LEARNING OBJECTIVES

- Name the functions that problem behavior can serve.
- Describe the role functional behavior assessment plays in preventing problem behavior and developing interventions for problem behavior.
- State the different methods of conducting a functional behavior assessment.
- State the primary characteristics of and rationale for conducting a descriptive assessment.
- Describe various methods for gathering descriptive assessment data and under what circumstances each is appropriate.
- Given a set of descriptive data, interpret the data to form a hypothesis regarding the possible function of problem behavior.
- State the primary characteristics of and rationale for conducting a functional analysis as a form of functional behavior assessment.
- Describe how to conduct a functional analysis.
- Given a set of data from a functional analysis, interpret the data to determine the function of problem behavior.

When it is time to wash hands before lunch, Stella turns the handles of the faucet and places her hands under the running water, but Flo screams and tantrums. The teacher is confused about why Stella performs so well in this context and Flo performs so poorly. Consequently, she is at a loss as to how to respond when Flo screams and tantrums. The crying and tantrums could occur for a variety of reasons. To understand why they are occurring, the teacher could perform a functional behavior assessment of Flo's problem behaviors. This assessment may provide information valuable for creating an effective behavior intervention. Consistent with the scientific precept of determinism described in Chapter 1, behaviors—including problem behaviors—are lawfully related to other events in the environment. **Functional behavior assessment (FBA)** enables hypotheses about the relations among specific types of environmental events and behaviors. Specifically, FBA is designed to obtain information about the purposes (functions) a behavior serves for a person. This chapter describes the basis for FBA, its role in the intervention and prevention of behavior difficulties, and alternative approaches to functional assessment. We will revisit Flo's problem behavior throughout this chapter to illustrate.

FUNCTIONS OF BEHAVIOR

Evidence from decades of research indicates that both desirable and undesirable behaviors, whether washing hands or screaming and tantrumming, are learned and maintained through interaction with the social and physical environment (see Schlinger & Normand, 2013, for a review). As explained in Chapters 11 and 12, these behavior–environment interactions are described as positive or negative reinforcement contingencies. Behaviors can be strengthened by either "getting something" or "getting out of something."

FBA is used to identify the type and source of reinforcement for challenging behaviors as the basis for intervention efforts designed to decrease the occurrence of those behaviors. FBA can be thought of as a reinforcer assessment of sorts. It identifies the reinforcers currently maintaining problem behavior. Those reinforcers might be positive or negative social reinforcers provided by someone who interacts with the person, or automatic reinforcers produced directly by the behavior itself. The idea behind FBA is that if these reinforcement contingencies can be identified, then interventions can be designed to decrease problem behavior and increase adaptive behavior by altering these contingencies. FBA fosters proactive, positive interventions for problem behavior. Although reinforcement contingencies are discussed in other chapters, a brief review of their role in FBA is warranted.

Positive Reinforcement

Social Positive Reinforcement (Attention)

Problem behavior often results in immediate attention from others, such as head turns; surprised facial expressions; reprimands; attempts to soothe, counsel, or distract; and so on. These reactions can positively reinforce problem behavior (even if

inadvertently), and the problem behavior is then more likely to occur in similar circumstances. Problem behavior maintained by positive reinforcement in the form of reactions from others can often occur in situations in which attention is otherwise infrequent, whether because the person does not have a repertoire to gain attention in desirable ways or because others in the environment are typically otherwise occupied.

Tangible Reinforcement

Many behaviors result in access to reinforcing materials or other stimuli. Just as pressing a button on the television remote changes the channel to a desired television show, problem behaviors can produce reinforcing outcomes. A child may cry and tantrum until a favorite television show is turned on; stealing another child's candy produces access to the item taken. Problem behaviors may develop when they consistently produce a desired item or event. This often occurs because providing the item temporarily stops the problem behavior (e.g., tantrum), although it can have the inadvertent effect of making the problem behavior more probable in the future under similar circumstances.

Automatic Positive Reinforcement

Some behaviors do not depend on the action of others to provide an outcome; some behaviors directly produce their own reinforcement. For example, thumb sucking might be reinforced by physical stimulation of either the hand or the mouth. Hand flapping while looking at a light might be reinforced by the visual stimulation of the light being blocked and revealed, as the hand flapping might simulate a flashing light. Swinging a tennis racket correctly as a tennis ball approaches and hits the head of the racket might be reinforced by the "pop" sound the ball makes when it hits the "sweet spot" of the racket. Placing hands under the faucet might be reinforced by the warm sensation of water on the skin. A behavior is assumed to be maintained by automatic reinforcement only after social reinforcers have been ruled out (e.g., when the behavior occurs even when the individual is alone).

Negative Reinforcement

Social Negative Reinforcement (Escape)

Many behaviors are learned as a result of their effectiveness in terminating or postponing aversive events. Hanging up the phone terminates interactions with a telemarketer; completing a task or chore terminates requests from others to complete it or the demands associated with the task itself. Problem behaviors can be maintained in the same way. Behaviors such as aggression, self-injurious behavior (SIB), and bizarre speech may terminate or avoid unwanted interactions with others. For example, noncompliance postpones engagement in a nonpreferred activity, and disruptive classroom behavior often results in the student being sent out of the classroom, thereby allowing escape from instructional tasks or teacher demands. All of these behaviors can be strengthened by negative reinforcement to the extent that they enable the individual to escape or avoid difficult or unpleasant tasks, activities, or interactions.

Automatic Negative Reinforcement

Aversive stimulation, such as a physically painful or uncomfortable condition, is a motivating operation that makes its termination reinforcing. Behaviors that directly terminate aversive stimulation are therefore maintained by negative reinforcement that is an automatic outcome of the response. Automatic negative reinforcement can account for behaviors that are either appropriate or harmful. For example, putting calamine lotion on a poison ivy rash can be negatively reinforced by alleviation of itching, but intense or prolonged scratching that breaks the skin can be negatively reinforced in the same manner. Some forms of SIB may distract from other sources of pain, which may account for their correlation with specific medical conditions (e.g., DeLissovoy, 1963).

Function Versus Topography

Several points can be made from the previous discussion of the sources of reinforcement for behavior. It is important to recognize that environmental influences do not make distinctions between desirable and undesirable topographies of behavior; the same reinforcement contingencies that account for desirable behavior can also account for undesirable behavior. For example, Stella, the child who washes and dries her hands before lunch, has probably received praise for doing so. Flo, the child who frequently engages in tantrums in the same situation, may have received attention (in the form of reprimands). Both forms of attention have the potential to reinforce the respective behaviors.

Likewise, the same topography of behavior can serve different functions for different individuals. For example, tantrums may be maintained by positive reinforcement in the form of attention for one child and by negative reinforcement in the form of escape for another child (e.g., Kennedy, Meyer, Knowles, & Shukla, 2000).

Because different behaviors that look quite different can serve the same function, and behavior of the same form can serve different functions under different conditions, the *topography*, or form, of a behavior often reveals little useful information about the conditions that account for it. Identifying the *conditions* that account for a behavior (its function), in contrast, suggests what conditions need to be altered to change the behavior. Assessing the function of a behavior can therefore yield useful information with respect to intervention strategies that are likely to be effective.

ROLE OF FUNCTIONAL BEHAVIOR ASSESSMENT IN INTERVENTION AND PREVENTION

FBA and Intervention

If a causal relation between environmental events and a problem behavior can be determined, and that relation can be altered, the problem behavior will occur less often. Interventions informed by FBA primarily consist of three strategic approaches: alter antecedent variables, alter consequent variables, and teach alternative behaviors.

Alter Antecedent Variables

FBA can identify antecedents that might be altered so the problem behavior is less likely to occur. Altering the antecedents

for problem behavior can change and/or eliminate either (a) the motivating operation for problem behavior or (b) the discriminative stimuli that trigger problem behavior. For example, the motivating operation for tantrums when Flo is asked to wash her hands before lunch could be modified by changing the characteristics associated with lunch so that the avoidance of particular events is no longer reinforcing. Depending on the function of the tantrums, this might involve initially reducing table-setting demands, altering seating arrangements to minimize taunts from a sibling or peer, or reducing snacks before lunch and offering more preferred foods during lunch. Alternatively, if the FBA shows that running water is the discriminative stimulus that triggers problem behavior when Flo is asked to wash her hands, she might be given a waterless antibacterial hand gel instead. In this case, the discriminative stimulus for problem behavior has been removed, thereby decreasing problem behavior.

Alter Consequent Variables

FBA can also identify a source of reinforcement to be eliminated for the problem behavior. For example, an FBA that indicates Flo's tantrums are maintained by social negative reinforcement (avoidance or escape) suggests a variety of treatment options, which, by altering that relation, are likely to be effective. Flo's tantrums can be placed on extinction by ensuring that the reinforcer (e.g., avoidance of the activities associated with setting the table for snack or lunch) no longer follows problem behavior (tantrums). Another approach could be to provide the reinforcer for alternative behaviors instead of for the problem behavior. Finally, the schedule might be modified so that hand washing follows—and thereby provides escape from—an event that is even less preferred.

Teach Alternative Behaviors

FBA can also identify the source of reinforcement to be provided for appropriate replacement behaviors. Alternative appropriate behaviors that have the same function (i.e., produce the same reinforcer) as tantrums could be taught. For example, if Flo's tantrums are maintained by escaping lunchtime activities, she might be taught to touch a card communicating "later" after washing her hands, to produce a delay in being seated at the lunch table.

FBA and Default Technologies

Interventions based on an FBA may be more effective than those selected without consideration of behavioral function (e.g., Ervin et al., 2001; Iwata et al., 1994b). Understanding *why* a behavior occurs often suggests *how* it can be changed for the better. In contrast, premature efforts to treat problem behavior before seeking an understanding of the purposes it has for a person can be inefficient, ineffective, and, in some cases, harmful. For example, suppose that a time-out procedure is implemented in an attempt to attenuate Flo's tantrums when she is asked to wash her hands before lunch. Flo is removed from the hand-washing activity to a chair in the corner of the room. It may be, however, that the events typically following hand washing (those associated with lunch time, such as the demands of arranging chairs or setting the table or interactions with others) are aversive for Flo, and tantrumming has allowed her to avoid those events. In this case, the intervention would be ineffective because it has done nothing to alter the relation between tantrums and the consequence of avoiding the aversive events associated with lunch. In fact, the intervention may exacerbate the problem if it produces a desired outcome for Flo. If stopping the hand-washing activity and having Flo sit on a chair as "time-out" for tantrumming enables her to avoid the aversive lunchtime events—or to escape them altogether—tantrums may be more likely under similar circumstances in the future. When the time-out intervention proves unsuccessful, other interventions might be attempted. Without understanding the function that the problem behavior serves, however, the effectiveness of those interventions cannot be predicted.

At best, a trial-and-error process of selecting and implementing interventions without consideration of behavioral function can be lengthy and inefficient. At worst, such an approach may cause the problem behavior to become more frequent or severe. As a result, caregivers might resort to increasingly intrusive, coercive, or punishment-based interventions, which are often referred to as *default technologies*.

FBA can decrease reliance on default technologies and contribute to more effective interventions in several ways. When FBAs are conducted, reinforcement-based interventions are more likely to be implemented than are those that include punishment (Pelios, Morren, Tesch, & Axelrod, 1999). In addition, the effects of interventions based on FBAs are likely to be more durable than those that do not take the function of problem behavior into account. If contrived contingencies are superimposed on unknown contingencies that are maintaining the behavior, their continuation is often necessary to maintain improvements in the behavior. If those superimposed contingencies are discontinued, the behavior will continue to be influenced by the unchanged operative contingencies.

FBA and Prevention

By furthering understanding of the conditions under which certain behaviors occur, FBA can also contribute to the prevention of difficulties. Although problem behavior may be suppressed without regard to its function by using punishment procedures, additional behaviors not subject to the punishment contingencies may emerge because the motivating operations for problem behavior remain. For example, contingent loss of privileges might eliminate tantrums that occur whenever Flo is asked to wash her hands, but it will not eliminate avoidance as a reinforcer or the conditions that establish avoidance as a reinforcer. Thus, other behaviors that result in avoidance, such as aggression, property destruction, or running away may develop. These unintended effects are less likely to occur with interventions that address (rather than override or compete with) the reinforcing functions of problem behavior.

On a broader scale, the accumulation of FBA data may further assist in prevention efforts by identifying the conditions that pose risks for the future development of problem behaviors. Preventive efforts can then focus on those conditions. For example, based on data from 152 analyses of the reinforcing functions of self-injurious behavior (SIB), Iwata and colleagues (1994b) found that escape from task demands or other aversive stimuli accounted for the behavior in the largest proportion of cases. The authors speculated that this outcome might have

been an unintended result of a move toward providing more aggressive treatment. For example, if Flo tantrums when she is required to wash her hands, the teacher might assume that she does not know how to wash her hands. The teacher might decide to replace playtime with a period of intensive instruction on hygiene. Rather than decreasing problem behavior, such interventions may exacerbate it. The data reported by Iwata and colleagues suggest that preventive efforts should be directed toward modifying instructional environments (such as providing more frequent reinforcement for desirable behavior, opportunities for breaks, or means to request and obtain help with difficult tasks) so that they are less likely to function as sources of aversive stimulation (motivating operations) for escape.

OVERVIEW OF FBA METHODS

FBA methods can be classified into three types: (a) functional (experimental) analysis, (b) descriptive assessment, and (c) indirect assessment. The methods can be arranged on a continuum with respect to considerations such as ease of use and the type and precision of information they yield (see Figure 27.1). Selecting the method or combination of methods that will best suit a particular situation requires consideration of each method's advantages and limitations. We discuss functional analysis and its variations first because descriptive and indirect methods of functional assessment developed as an outgrowth of functional analysis. As noted later, functional analysis is the only FBA method that allows practitioners to confirm hypotheses regarding functional relations between problem behavior and environmental events.

Functional (Experimental) Analysis

Basic Procedure

Functional analysis procedures were first pioneered by Iwata, Dorsey, Slifer, Bauman, and Richman (1994a). In a **functional analysis**, antecedents and consequences representing those in the person's natural environment are arranged so that their separate effects on problem behavior can be observed and measured. This type of assessment is often referred to as an *analog* because antecedents and consequences similar to those occurring in the natural routines are presented in a systematic manner, but the analysis is not conducted in the context of naturally occurring routines. Analog conditions are often used because they allow the behavior analyst to better control the environmental variables that may be related to the problem behavior than can be accomplished in naturally occurring situations. Analogs refer to the arrangement of variables rather than the setting in which assessment occurs. Research has found that functional analyses conducted in natural environments (e.g., classroom settings) often yield the same (and, in some cases, clearer) results compared to those conducted in simulated settings (Noell, VanDerHeyden, Gatti, & Whitmarsh, 2001).

Functional analyses (FA) typically comprise four conditions: three test conditions—contingent attention, contingent escape, and alone—and a control condition, in which problem behavior is expected to be low because reinforcement is freely available and no demands are placed on the individual (see Table 27.1). However, it should be understood that there is no "standard" FA. FAs should be flexible and individualized. The practitioner can implement myriad conditions to address

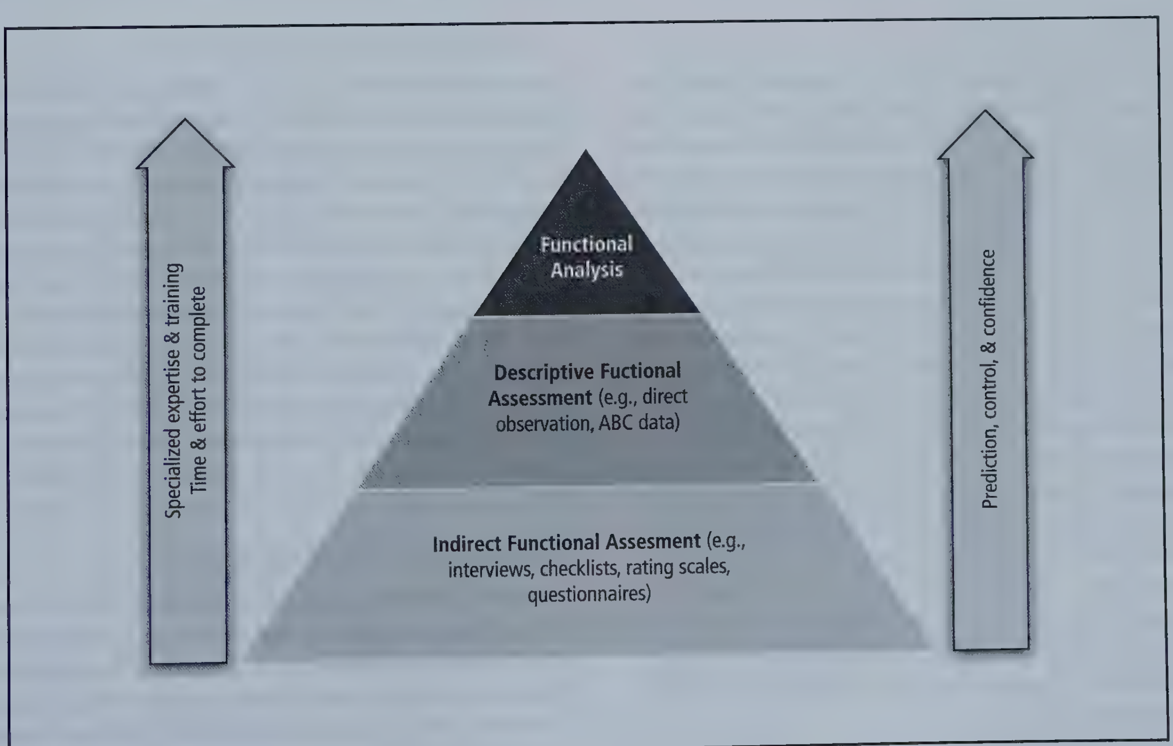


Figure 27.1 FBA methods. The amount or value of each dimension or outcome identified within the arrows increases from the bottom to the top of the pyramid.

Diagram contributed by Rebecca Eldridge.

TABLE 27.1 Motivating Operations and Reinforcement Contingencies for Typical Control and Test Conditions of a Functional Analysis

Condition	Antecedent conditions (motivating operation)	Consequences for problem behavior
Play (control)	Preferred activities are continuously available, social attention is provided, and no demands are placed on the person.	Problem behavior is ignored or neutrally redirected.
Contingent attention	Attention is diverted or withheld from the person.	Attention is provided in the form of mild reprimands or soothing statements (e.g., "Don't do that. You'll hurt someone.").
Contingent escape	Task demands are delivered continuously using a three-step prompting procedure (e.g., [1] "You need to fold the towel." [2] Model folding the towel. [3] Provide hand-over-hand assistance to fold the towel).	Break from the task is provided by removing task materials and stopping prompts to complete the task.
Contingent tangible	After having access to a preferred activity/toy, the preferred activity/toy is removed (e.g., "My turn to play with that."). Moderately preferred or neutral toys are still continuously available, as is adult attention.	The preferred activity/toy is returned to the individual.
Alone	Low level of environmental stimulation is present (i.e., therapist, task materials, and play materials are absent).	Problem behavior is ignored or neutrally redirected.

specific hypotheses. For example, some FAs also include a tangible test condition (Day, Rea, Schussler, Larsen, & Johnson, 1988). Each test condition contains a motivating operation (MO) and a potential source of reinforcement for problem behavior. The conditions are presented systematically one at a time and in an alternating sequence, and occurrences of problem behavior are recorded during each session. Sessions are repeated to determine the extent to which problem behavior consistently occurs more often under one or more conditions relative to another.

Interpreting Functional Analyses

The function problem behavior serves for a person can be determined by visually inspecting a graph of the results of an analysis to identify the condition(s) under which high rates of the behavior occurred. A graph for each potential behavioral function is shown in Figure 27.2. Problem behavior is expected to be low in the play condition because no motivating operations for problem behavior are present. Elevated problem behavior in the contingent attention condition suggests that problem behavior is maintained by social positive reinforcement (see graph in top left of Figure 27.2). Elevated problem behavior in the contingent escape condition suggests that problem behavior is maintained by negative reinforcement (graph in top right of Figure 27.2). Elevated problem behavior in the alone condition suggests that problem behavior is maintained by automatic reinforcement (graph in bottom left of Figure 27.2). Further analysis is needed to determine if the source of the automatic reinforcement is positive or negative. Problem behavior may be maintained by multiple sources of reinforcement. For example, if problem behavior is elevated in the contingent attention and contingent escape conditions, it is most likely maintained by both positive and negative reinforcement.

If problem behavior occurs frequently in all conditions (including the play condition), or is variable across conditions, responding is considered *undifferentiated* (see graph in bottom right of Figure 27.2). Such results are inconclusive, but can also occur with behavior that is maintained by automatic reinforcement.

Functional analysis has been replicated and extended in hundreds of studies, thereby demonstrating its generality as an approach to the assessment and treatment of a wide range of behavior difficulties. (See the 1994 and 2013 special issues of the *Journal of Applied Behavior Analysis* for a sample of such applications and Hanley, Iwata, & McCord, 2003, and Beavers, Iwata, & Lerman, 2013, for reviews.)

Advantages of Functional Analysis

The primary advantage of functional analysis is its ability to yield a clear demonstration of the variable or variables that influence the occurrence of a problem behavior. In fact, functional (experimental) analyses are the standard of scientific evidence by which other assessment alternatives are evaluated, and represent the method most often used in research on the assessment and treatment of problem behavior (Arndorfer & Miltenberger, 1993). Because functional analyses allow valid conclusions concerning the variables that maintain problem behavior, they have enabled the development of effective reinforcement-based treatments and less reliance on punishment procedures (Brosnan & Healy, 2011; Ervin et al., 2001; Horner, Carr, Strain, Todd, & Reed, 2002; Iwata et al., 1994b; Pelios et al., 1999).

Limitations of Functional Analysis

A risk of functional analysis, however, is that the assessment process may temporarily strengthen or increase the undesirable behavior to unacceptable levels, or possibly result in the behavior acquiring new functions. Second, relatively little is known about the acceptability of functional analysis procedures to practitioners (Ervin et al., 2001; Reid & Nelson, 2002). The deliberate arrangement of conditions that set the occasion for, or potentially reinforce, problem behavior can

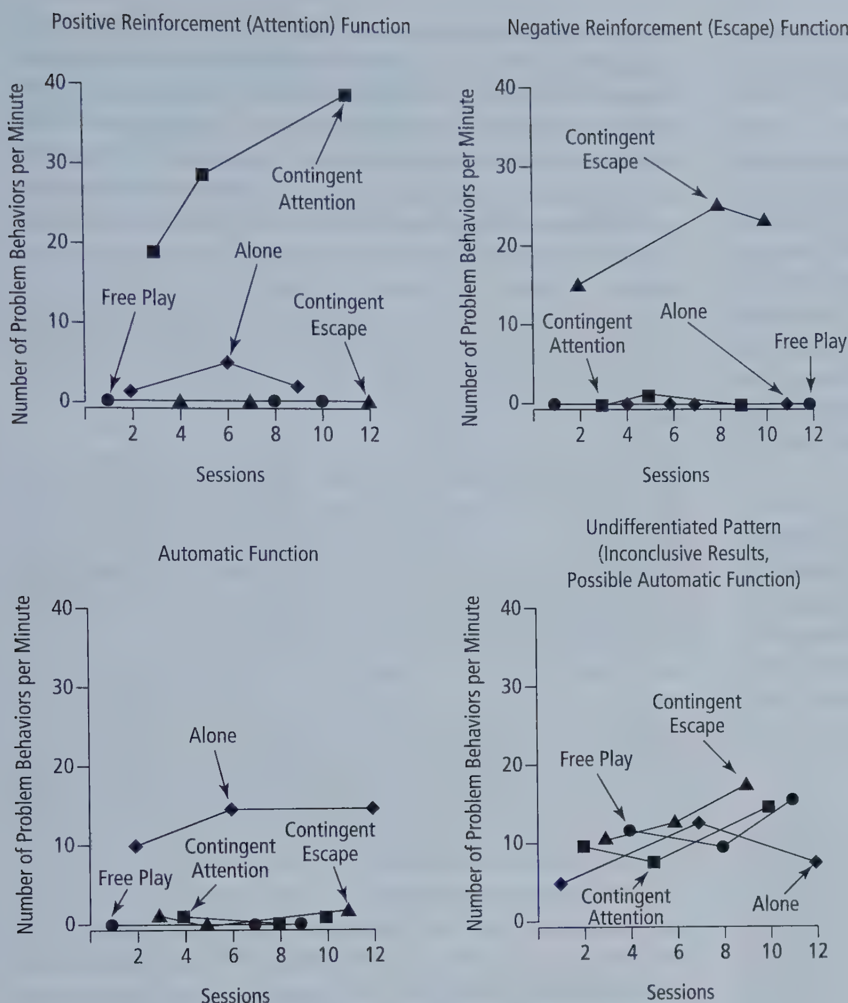


Figure 27.2 Data patterns typical of each behavioral function during a functional analysis.

be counterintuitive to persons who do not understand its purpose (or that such conditions are analogs for what occurs in the natural routine). O'Neill, Bundock, Kladis, and Hawken (2015) found that teachers reported positive perceptions of, and willingness to engage in, systematic direct observations and functional analysis manipulations. Interestingly, their positive perceptions were reportedly higher than those of school psychologists. School psychologists also expressed more concern about the time and effort required to conduct FBAs. Across both teachers and school psychologists, the most concerns were expressed about FA procedures. O'Neill and colleagues suggest that some of the concerns about these issues emanated from a lack of training and a lack of time to conduct the analyses. A third limitation is that some behaviors (e.g., those that, albeit serious, occur infrequently) may not be amenable to functional analyses due to setting and other factors. For example, if proper safety equipment or individuals with enough training to keep the client safe are not available, the setting may not be safe for conducting the FA. Fourth, functional analyses that are conducted in contrived settings might not detect the variables that account for the occurrence of the problem behavior in the natural environment, particularly if the behavior is controlled by idiosyncratic variables that are not represented in the functional analysis conditions (e.g., Noell et al., 2001). Finally, the time, effort, and professional expertise required to conduct and interpret functional analyses have been frequently cited as obstacles to its widespread use in practice (e.g., Spreat & Connelly, 1996; 2001, Volume 2 issue of *School Psychology Review*), although Hanley (2012) argues that these perceptions are not warranted.

Of course, untreated or ineffectively treated problem behaviors also consume a great deal of time and effort (without a constructive long-term outcome), and implementation of effective treatments (based on an understanding of the variables that maintain them) is likely to require skills similar to those involved in conducting functional analyses. These concerns have led to research on ways to enhance the practical use of functional analyses, including variations and the development of alternative methods of FBA described in the following sections.

Variations of Functional Analysis Procedures

Rather than dictating particular settings or conditions, functional analysis procedures achieve relevance through flexible arrangements tailored to individual circumstances. Some procedural variations are described here; for a review, see Lydon, Healy, O'Reilly, and Lang (2012).

Brief Functional Analysis. Functional analyses are typically conducted across multiple sessions. Sometimes time constraints or the severity of the problem behavior may restrict the number of sessions that can be conducted. In such circumstances, a **brief functional analysis** (Derby et al., 1992; Northup et al., 1991) may be appropriate. In a brief functional analysis, only one or two 5- to 10-min sessions are conducted for each condition. A convincing demonstration of function can be achieved by either alternating a condition that produces problem behavior with one that does not or conducting a contingency reversal. A contingency reversal is when the practitioner first makes the putative reinforcer contingent on one target behavior, such as tantrums, and then on an appropriate replacement behavior, such as making requests. (See the case example with Marie later in this chapter for an illustration.)

A brief functional analysis can be a robust procedure (Tincani, Castrogiavanni, & Axelrod, 1999) that preserves many advantages of an FA; however, it reveals a function clearly in fewer cases than a full functional analysis (Kahng & Iwata, 1999; Wallace & Knights, 2003). Some researchers (e.g., Vollmer, Marcus, Ringdahl, & Roane, 1995) have suggested that practitioners may want to consider beginning with a brief functional analysis and proceeding to a full functional analysis only if clear and replicable patterns do not emerge.

Conducting Functional Analyses in Natural Settings. Although not a requirement of functional analyses, they are often conducted in clinical settings. However, it can be advantageous to conduct functional analyses in the setting where the problem behaviors typically occur because the relevant stimulus cues are present in those settings. Sasso and colleagues (1992) provided one of the first demonstrations of functional analysis in classroom settings. Since then, many others have demonstrated the utility of functional analysis in the classroom (e.g., Wright-Gallo, Higbee, Reagon, & Davey, 2006), vocational (e.g., Wallace & Knights, 2003), and home settings (e.g., Wacker et al., 2013). One challenge when implementing functional analyses in natural settings is that the procedures can be disruptive to ongoing routines and may require additional supervision so that the other students, children, or clients are appropriately supervised while the teacher or caregiver devotes his or her attention to the analysis. An adaptation that can be made in natural settings is to use a trial-based approach, which may be more easily embedded into the ongoing activities of the classroom (Sigafoos & Saggers, 1995).

Trial-based Functional Analysis. A **trial-based functional analysis** (Bloom, Iwata, Fritz, Roscoe, & Carreau, 2011; Lambert, Bloom, & Irvin, 2012; Sigafoos & Sagers, 1995) consists of a series of trials interspersed among classroom activities. (See the case example with Carson later in this chapter for an illustration.) Each trial consists of two components, each lasting 1 minute. The first component consists of presenting the establishing operation and contingency for problem behavior (test condition), and the second component consists of continuous access to the reinforcer (control condition). For example, during the first component, a demand may be placed on the child and as soon as a tantrum occurs, the demand is terminated. During the second component, a task break is continuously provided. These conditions are implemented at times when they naturally occur (e.g., during free play periods for tangible conditions and during classroom instruction for escape conditions). Teachers may be able to perform the procedure with fidelity in their classrooms (Rispoli et al., 2015) and find it more manageable than session-based functional analyses. However, Bloom et al. describe a number of considerations that may result in discrepant results using trial-based functional analyses. For example, exposure to the relevant establishing operations and consequences may be too brief to evoke or maintain problem behavior.

Synthesized Functional Analyses. Hanley, Jin, Vanselow, and Hanratty (2014) and Jessel, Hanley, and Ghaemmaghami (2016) describe a variation of functional analysis that is designed to increase efficiency by using an **interview-informed synthesized contingency analysis** (IISCA). In the test condition, multiple contingencies are implemented simultaneously (e.g., attention and escape) when the problem behavior is demonstrated. In the control condition, those same reinforcers are presented noncontingently and continuously. For example, for one participant, Gail, in Hanley et al. (2014), problem behavior was not observed in the typical attention and tangible conditions, even though the behavior analysts hypothesized these were the variables maintaining her problem behavior. When these two contingencies were presented together (i.e., both attention *and* tangibles presented simultaneously contingent on problem behavior), a clear pattern of problem behavior emerged, as compared to a condition in which attention and tangibles were available continuously. Furthermore, these effects were observed only when Gail's mother presented the contingencies and not when the behavior analyst presented them.

This technology may have particular utility when problem behavior is not observed in typical FA conditions and when the behavior analyst suspects multiple contingencies may be maintaining problem behavior. At this time, the limited research that has compared IISCA with functional analysis is conflicting (Fisher, Greer, Romani, Zangrillo, & Owen, 2016; Hanley et al., 2014). Thus, the utility of IISCA is as yet unclear. We suggest that behavior analysts closely monitor additional research on this topic as it emerges.

Latency-based Functional Analysis. Sometimes, problem behavior is so severe that it is not possible to conduct an assessment that requires repeated occurrences of the problem behavior. In these cases, a **latency-based functional analysis** may be helpful (Thomasen-Sassi, Iwata, Neidert, & Roscoe, 2011). (See the case example with Elija later in this chapter for an illustration.) In a latency-based functional analysis, like a trial-based assessment, each session is terminated as soon as a problem behavior occurs. However, in this case, the session may last longer than 1 minute, because the establishing operation is present as long as necessary to evoke the problem behavior (or until a pre-established time limit has expired). The index of problem behavior is the latency from onset of the

establishing operation to the first occurrence of the problem behavior. Using a latency-based measure can reduce the number of occurrences of problem behavior necessary for analysis and shorten the time required to conduct the analysis. As with trial-based assessments, a limitation to latency-based assessment is that there is only one exposure to the contingency per session. Analyses that rely on response-repetition measures are generally preferable.

Functional Analysis of Precursors. Sometimes, even one occurrence of problem behavior can be so severe that evoking it poses unacceptable risk to the client and/or others. In these cases, it can be helpful to conduct the functional analysis on precursor behavior—a behavior that reliably precedes the target behavior (Borrero & Borrero, 2008; Herscovitch, Roscoe, Libby, Bourret, & Ahearn, 2009; Najdowski, Wallace, Ellsworth, MacAleese, & Cleveland, 2008; Smith & Churchill, 2002). (See the case example with Will later in this chapter for an illustration.) For example, growling may reliably signal aggression, in which case the analysis can be conducted using growling.

In addition to these variations of functional analysis, other methods of FBA can be used to inform the functional analysis process. Although these methods have sometimes been used as alternatives to functional analyses, we view them as valuable steps in the process, not as substitutes for a functional analysis.

Safety Considerations for FAs

When conducting FAs, the behavior analytic practitioner must consider ethical issues surrounding the client's safety. Kahng et al. (2015) noted that, in a retrospective study of 99 clients' functional analyses, the injury rate was relatively low and any injuries that did occur were mild. However, because target behaviors are expected to be evoked during FAs, it is important to implement precautionary measures to ensure the safety of all individuals involved. In particular, steps should be taken when analyzing potentially harmful behaviors, such as aggression, self-injurious behavior, and pica. Some procedures that have been used to mitigate risk of injury when assessing potentially dangerous behaviors are as follows: medical examinations to rule out health and physical concerns, presence of medical staff for especially severe behavior (Iwata 1982, 1994a), ensuring that trained practitioners plan and run analyses, having additional staff present to block behavior, using protective equipment, establishing a behavioral and/or medical criterion for terminating sessions (Kahng et al., 2015), using soft toys to conduct the analysis (Hanley, 2012), and using brief FA or assessing precursor behaviors (Najdowski et al., 2008; Smith & Churchill, 2002). Protective equipment can be utilized during the analysis as well; however, researchers have found the use of protective equipment can invalidate FA results (Borrero, Vollmer, Wright, Lerman, & Kelley, 2002; Le & Smith, 2002). (For a list of some other protections that can be taken, see Hanley, 2012, and Weeden, Mahoney, & Poling, 2010). Thus, although these precautions are often cited as recommendations, additional research is warranted to develop widely accepted evidence-based safety standards and best-practice guidelines.

Descriptive Functional Behavior Assessment

As with functional analyses, **descriptive functional behavior assessment** encompasses direct observation of behavior; unlike functional analyses, however, observations are made under naturally occurring conditions. Thus, descriptive assessments involve observation of the problem behavior in relation to events that are not arranged in a systematic manner. Descriptive assessments have roots in the early stages of applied behavior analysis; Bijou, Peterson, and Ault (1968) initially described a method for objectively defining, observing, and coding behavior and contiguous environmental events. This method has been used subsequently to identify events that may be correlated with the target behavior. Events that are shown to have a high degree of correlation with the target behavior may suggest hypotheses about behavioral function. We describe three variations of descriptive analysis: ABC (antecedent-behavior-consequence) continuous recording, ABC narrative recording, and scatterplots.

Before we describe the forms of descriptive assessment, we think it is important for the reader to consider some caveats. All of the forms of descriptive assessments explained below are generally considered invalid for detecting behavioral function. This is because they tend to yield false positives for an attention function. When an individual displays severe problem behavior, attention is often ubiquitous in the environment—attention is often delivered both independent of and contingent on problem behavior. In addition, false negatives can occur for escape-maintained problem behavior, because others in the environment may deliberately avoid making demands on an individual or placing the individual in any situations that trigger the problem behavior. Instead, the reader is cautioned to restrict conclusions based on descriptive assessments to only the prevalence of environmental factors that precede and follow the target behavior (Hanley, 2012).

ABC Continuous Recording

With ABC continuous recording, an observer records occurrences of the targeted problem behaviors and selected environmental events in the natural routine during a period of time. Codes for recording specific antecedents, problem behaviors, and consequences can be developed based on information obtained from a Functional Assessment Interview or ABC narrative recording (described later). For example, following an interview and observations using narrative recording, Lalli, Browder, Mace, and Brown (1993) developed stimulus and response codes to record the occurrence or nonoccurrence of antecedent (e.g., one-to-one instruction, group instruction) and subsequent events (attention, tangible reinforcement, escape) for problem behavior during classroom activities.

With ABC continuous recording, the occurrence of a specified event is marked on the data sheet (using partial interval, momentary time sampling, or frequency recording)

(see Figure 27.3). The targeted environmental events (antecedents and consequences) are recorded whenever they occur, regardless of whether problem behavior occurred with it. Recording data in this manner may reveal events that occur in close temporal proximity to the target behavior. For example, descriptive data may show that tantrums (behavior) often occur when a student is given an instruction to wash her hands (antecedent); the data may also show that tantrums are typically followed by the removal of task demands. A possible hypothesis in this situation is that tantrums are motivated by academic demands and are maintained by escape from those demands (negative reinforcement). (See the case example with Chris later in this chapter for an illustration.)

Advantages of ABC Continuous Recording. Descriptive assessments based on continuous recording use precise measures (similar to functional analyses), and in some cases the correlations may reflect causal relations (e.g., Alter, Conroy, Mancil, & Haydon, 2008; Sasso et al., 1992). Because the assessments are conducted in the context in which the problem

Figure 27.3 Sample data collection form for ABC continuous recording.

ABC Recording Form
observer: R. Van Norman
Time begin: 9.30 A.M Time end: 10:15 A.M.
Date: January 25, 2006

Antecedent	Behavior	Consequence
☐ Task prompt/instruction ☒ Attention diverted ☐ Social interaction ☐ Engaged in preferred activity ☐ Preferred activity removed ☐ Alone (no attention/no activities)	☒ Tantrum ☐ Aggression	☐ Social attention ☒ Reprimand ☐ Task demand ☐ Access to preferred item ☐ Task removed ☐ Attention diverted
☒ Task prompt/instruction ☐ Attention diverted ☐ Social interaction ☐ Engaged in preferred activity ☐ Preferred activity removed ☐ Alone (no attention/no activities)	☒ Tantrum ☐ Aggression	☐ Social attention ☐ Reprimand ☐ Task demand ☐ Access to preferred item ☒ Task removed ☐ Attention diverted
☒ Task prompt/instruction ☐ Attention diverted ☐ Social interaction ☐ Engaged in preferred activity ☐ Preferred activity removed ☐ Alone (no attention/no activities)	☒ Tantrum ☐ Aggression	☐ Social attention ☐ Reprimand ☐ Task demand ☐ Access to preferred item ☒ Task removed ☐ Attention diverted
☐ Task prompt/instruction ☒ Attention diverted ☐ Social interaction ☐ Engaged in preferred activity ☐ Preferred activity removed ☐ Alone (no attention/no activities)	☒ Tantrum ☐ Aggression	☐ Social attention ☐ Reprimand ☐ Task demand ☐ Access to preferred item ☐ Task removed ☒ Attention diverted
☒ Task prompt/instruction ☐ Attention diverted ☐ Social interaction ☐ Engaged in preferred activity ☐ Preferred activity removed ☐ Alone (no attention/no activities)	☐ Tantrum ☐ Aggression	☐ Social attention ☐ Reprimand ☐ Task demand ☐ Access to preferred item ☐ Task removed ☐ Attention diverted
☒ Task prompt/instruction ☐ Attention diverted ☐ Social interaction ☐ Engaged in preferred activity ☐ Preferred activity removed ☐ Alone (no attention/no activities)	☒ Tantrum ☐ Aggression	☐ Social attention ☐ Reprimand ☐ Task demand ☐ Access to preferred item ☒ Task removed ☐ Attention diverted

(continued)

Figure 27.3 (*continued*)

☐ Task prompt/instruction ☒ Attention diverted ☐ Social interaction ☐ Engaged in preferred activity ☐ Preferred activity removed ☐ Alone (no attention/no activities)	☐ Tantrum ☐ Aggression	☐ Social attention ☐ Reprimand ☐ Task demand ☐ Access to preferred item ☐ Task removed ☐ Attention diverted
☒ Task prompt/instruction ☐ Attention diverted ☐ Social interaction ☐ Engaged in preferred activity ☐ Preferred activity removed ☐ Alone (no attention/no activities)	☒ Tantrum ☐ Aggression	☐ Social attention ☐ Reprimand ☐ Task demand ☐ Access to preferred item ☒ Task removed ☐ Attention diverted
☐ Task prompt/instruction ☐ Attention diverted ☒ Social interaction ☐ Engaged in preferred activity ☐ Preferred activity removed ☐ Alone (no attention/no activities)	☒ Tantrum ☐ Aggression	☐ Social attention ☒ Reprimand ☐ Task demand ☐ Access to preferred item ☐ Task removed ☐ Attention diverted
☒ Task prompt/instruction ☐ Attention diverted ☐ Social interaction ☐ Engaged in preferred activity ☐ Preferred activity removed ☐ Alone (no attention/no activities)	☒ Tantrum ☐ Aggression	☐ Social attention ☐ Reprimand ☐ Task demand ☐ Access to preferred item ☒ Task removed ☐ Attention diverted
☐ Task prompt/instruction ☐ Attention diverted ☒ Social interaction ☐ Engaged in preferred activity ☐ Preferred activity removed ☐ Alone (no attention/no activities)	☒ Tantrum ☐ Aggression	☐ Social attention ☐ Reprimand ☐ Task demand ☐ Access to preferred item ☐ Task removed ☒ Attention diverted

Source: Recording form developed by Renée Van Norman. Used by permission.

behaviors occur, they are likely to provide useful information for designing a subsequent functional analysis if that proves necessary. In addition, they do not require disruption to the person's routine.

Considerations for ABC Continuous Recording. Although descriptive analyses of this type may show a correlation between particular events and the problem behavior, such correlations can be difficult to detect in many situations. This is especially likely if the influential antecedents and consequences do not reliably precede and follow the behavior. In such cases, it may be necessary to analyze descriptive data by calculating **conditional probability** (the likelihood that a target problem behavior will occur in a given circumstance) or conducting a **contingency space analysis** (CSA, see Box 27.1).

However, conditional probabilities can be misleading. If a behavior is maintained by intermittent reinforcement, the behavior might occur often even though it is not followed consistently by a particular consequence. For example, the teacher might send the student to time-out only when tantrums are so frequent or severe that they are intolerable. In this case, only a small proportion of tantrums would be followed by that consequence, and the conditional probability would be low. One possibility, therefore, is that a functional relation that exists (e.g., tantrums negatively reinforced by escape) will not be detected. Furthermore, the child's current behavior intervention plan might require three repetitions of the instruction and an attempt to provide physical assistance before time-out is implemented. In that situation, the conditional probability of tantrums being followed by attention would be high. A descriptive analysis might, therefore, suggest a functional relation (e.g., tantrums positively reinforced by attention) that does *not* exist. Perhaps for these reasons, studies that have used conditional probability calculations to examine the extent to which descriptive methods lead to the same hypotheses as functional analyses have generally found low agreement (e.g., Lerman & Iwata, 1993; Noell et al., 2001).

ABC Narrative Recording

ABC narrative recording is a form of descriptive assessment that differs from continuous recording in that (a) data are collected only when behaviors of interest are observed, and (b) the

BOX 27.1

What Is Contingency Space Analysis?

by Nathan VanderWeele and Corinne Gist

Contingency spaces display the relationship between events in terms of the probability of one event (e.g., a particular consequence) given or not given another event (e.g., the presence or absence of a particular behavior) (Martens, DiGennaro, Reed, Szczech, & Rosenthal, 2008; Matthews, Shimoff, & Catania, 1987; Schwartz, 1989). For example, if Johnny yells out 13 times during 20 observation intervals and, as a result, the teacher removes him from the classroom following 2 of those occurrences, we would say that the probability of the teacher removing Johnny from the classroom was 2/13, or approximately 15% of occurrences (see raw data below in row, Escape/Present; columns B, B+C, and B+C/B).

Attention	B	B+C	B+C/B
Present	13	9	0.69
Absent	7	3	0.43
Escape	B	B+C	B+C/B
Present	13	2	0.15
Absent	7	3	0.43
Other	B	B+C	B+C/B
Present	13	2	0.14
Absent	7	1	0.17

However, identifying the probability of a behavior–consequence relation in isolation represents an incomplete analysis of the event. In addition to determining the probability of the occurrence of the consequence following the behavior, the probability of the consequence following the absence of the behavior should also be determined, (e.g., 43% per the raw data shown above in row, Escape/Absent; columns B, B+C, and B+C/B) with the resulting values compared in what is called a *contingency space analysis (CSA).*

Graphing conditional probabilities in the general operant contingency space reveals the degree of contingency between events and behavior (see Figure A below), which may address some of the limitations associated with ABC recordings. A CSA is conducted in the natural environment and does not require experimental manipulation of the consequences. Direct observation is used to identify the presence of a positive, negative, or neutral contingency between behavior and its consequences. The contingency is said to be "positive" if a specific consequence is more likely to occur after the behavior. The contingency is said to be "negative" if the consequence is less likely to occur following the behavior. For example, if Johnny yells out 13 times during an observation period and receives teacher attention following 9 of those instances, we would say that the consequence of teacher attention (c) was *more* likely to occur following yelling out (b), indicating a positive contingency for yelling out and teacher attention. Alternatively, if the teacher more frequently withholds attention (c) following yelling out (b), we would say that teacher attention is *less* likely to occur following the behavior. Thus, a negative contingency may exist, since the consequence is not likely to follow the behavior.

As previously mentioned, to determine the operant strength of a behavior–consequence relation, the probability of the consequence following the behavior *and* the

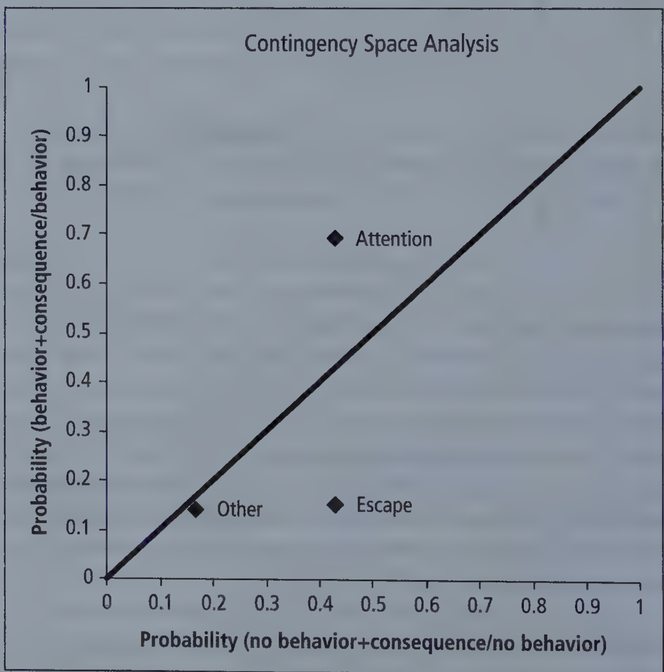


Figure A Hypothetical CSA data with graph illustrating potential positive (attention), negative (escape), and neutral (other) contingencies provides an example of a contingency space graph. The *y*-axis value equals the probability of a consequence given the occurrence of a behavior, and the *x*-axis value equals the probability of a consequence given the absence of a behavior.

probability of the consequence following the absence of behavior are calculated. To calculate the probability, the number of times the consequence followed the behavior is divided by the total number of behaviors emitted during an observation period, and the number of times the consequence occurred following the absence of the behavior is divided by the total number of intervals in which the behavior did not occur. These values are then plotted along the *x*-axis of the graph (see Figure A). The location of the resulting data point on the graph indicates the type and strength of the behavior–consequence relation. If the data point falls above and to the left of the line, a positive contingency may exist, suggesting that the consequence could be maintaining the behavior. If the data point falls below and to the right of the line, a negative contingency may exist, suggesting that the consequence is likely not maintaining the behavior. If data points fall on or near the neutrality line, no contingency is likely to exist, indicating that the consequence is equally likely to occur independent of the occurrence of behavior. Additionally, the greater the distance between the data point and the neutrality line, the stronger (positive) or weaker (negative) the hypothesized behavior–consequence relationship. It is important to note, however, that the efficacy of CSA results will rely heavily on the fidelity of the descriptive data collected.

recording is open-ended (any events that immediately precede and follow the target behavior are noted) (see Figure 3.3 in this text). Because data are recorded only when the target behavior occurs, narrative recording may be less time-consuming than continuous recording. However, narrative recording has several disadvantages in addition to those described earlier.

Considerations for Narrative Recording. Because narrative recording data are seldom reported in published research, their utility in identifying behavioral function has not been established. However, ABC narrative recording might identify functional relations that do not exist because antecedent and consequent events are recorded only in relation to the target behavior; it would not be evident from ABC data if particular events occurred just as often in the absence of the target behavior. For example, ABC data might erroneously indicate a correlation between peer attention and disruption even though peer attention also occurred frequently when the student was not disruptive.

Another potential limitation of ABC narrative recording concerns its accuracy. Unless observers receive adequate training, they may report inferred states or subjective impressions (e.g., "felt embarrassed," "was frustrated") instead of describing observable events in objective terms. In addition, given the likelihood that a number of environmental events occur in close temporal proximity to one another, discriminating the events that occasion a behavior can be difficult. Lerman, Hovanetz, Strobel, and Tetreault (2009) found that teachers collected ABC data more accurately when using a more structured format than with narrative recording. Thus, ABC narrative recording may be best suited as a means of gathering preliminary information to inform continuous recording or functional analyses.

Scatterplot Recording

Scatterplot recording is a procedure for recording the extent to which a target behavior occurs more often at particular times than others (Symons, McDonald, & Wehby, 1998; Touchette, MacDonald, & Langer, 1985). Specifically, scatterplots involve dividing the day into blocks of time (e.g., a series of 30-min segments). For every time segment, an observer uses different symbols on an observation form to indicate whether the target problem behavior occurred a lot, some, or not at all (see Figure 27.4). After data have been collected over a series of days, they are analyzed for patterns (specific time periods that are typically associated with problem behavior). If a recurring response pattern is identified, the temporal distributions in the behavior can be examined for a relation to particular environmental events. For example, a time period in which the behavior occurs often might be correlated with increased demands, low attention, certain activities, or the presence of a particular person. If so, changes can be made on that basis.

Considerations for Scatterplots. The primary advantage of scatterplots is that they identify time periods during which the problem behavior occurs. Such information can be useful in pinpointing periods of the day when more focused ABC assessments might be conducted to obtain additional information regarding the function of the problem behavior. Although scatterplots are often used in practice, little is known about their utility. It is unclear whether temporal patterns are routinely evident (Kahng et al., 1998). Another problem is that obtaining accurate data with scatterplots may be difficult (Kahng et al., 1998). In addition, the subjective nature of the ratings of how often the behavior occurs (e.g., "a lot" versus "some") can contribute to difficulties with interpretation (standards for these values might differ across teachers or raters).

Indirect Functional Behavior Assessment

Indirect functional assessment methods use structured interviews, checklists, rating scales, or questionnaires to obtain information from those who are familiar with the person exhibiting the problem behavior (e.g., teachers, parents, caregivers, and/or the individual himself or herself) to identify possible conditions or events in the natural environment that correlate with the

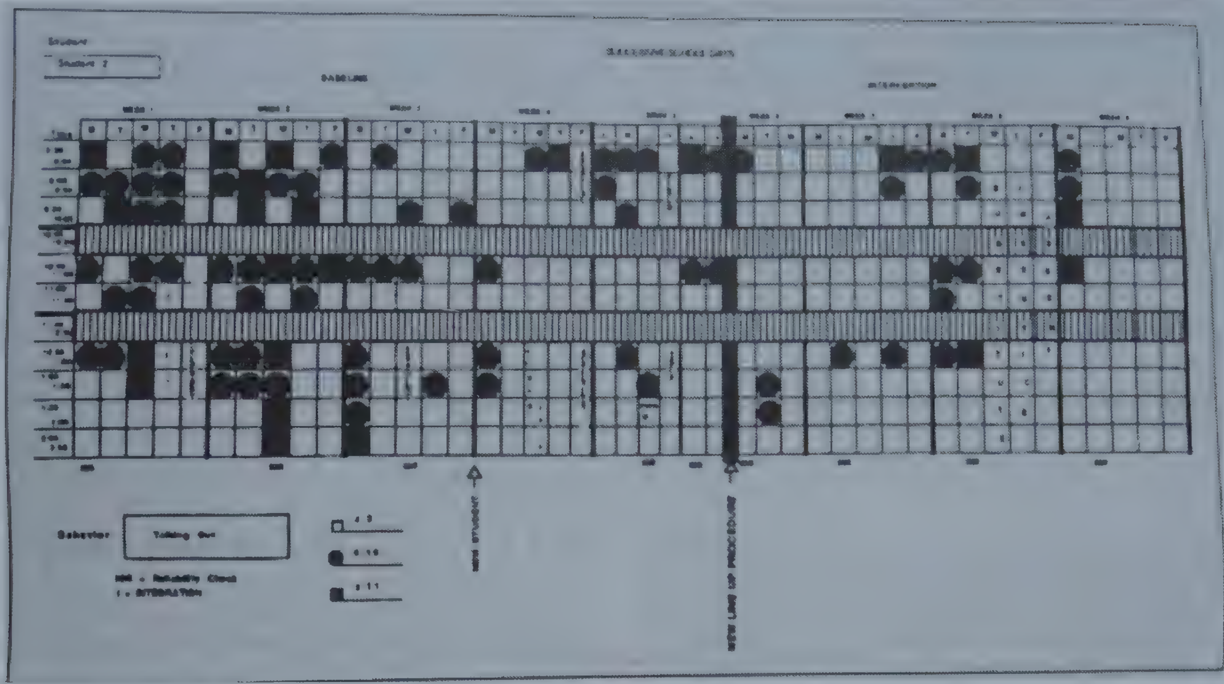


Figure 27.4 Example of scatterplot recording form. Different shapes represent different observed rates of talk-outs per daily observation intervals. Open squares, 2 or fewer talk-outs; shaded circles, 4–10 talk-outs; filled squares, 11 or more occurrences.

From "Functional Assessment and Teacher Collected Data," by F. J. Symons, L. M. McDonald, and J. H. Wehby, 1998, *Education and Treatment of Children*, *21*, p. 145. Copyright 1998 Used by permission.

problem behavior. Such procedures are referred to as "indirect" because they do not involve direct observation of the behavior, but rather solicit information based on others' recollections of the behavior.

Behavioral Interviews

Interviews are used routinely in assessment. The goal of a behavioral interview is to obtain clear and objective information about the problem behaviors, antecedents, and consequences. This might include clarifying descriptions of the behavior (consequences); when (times), where (settings, activities, events), with whom, and how often it occurs; what typically precedes the behavior (antecedents); what the child and others typically do immediately following the behavior (consequences); and what steps have previously been taken to address the problem, and with what result. Similar information might be solicited about desirable behavior (or the conditions under which undesirable behavior does not occur) to identify patterns or conditions that predict appropriate versus problem behavior. Information can also be obtained about the child's apparent preferences (e.g., favorite items or activities), skills, and means of communicating. A skillful interviewer poses questions in a way that evokes specific, complete, and factual responses about events, with minimal interpretation or inferences.

Lists of interview questions have been published, and they provide a consistent, structured format for obtaining information through either an interview or a questionnaire format. For example, the Functional Assessment Interview (O'Neill et al., 1997) has 11 sections, which include description of the form (topography) of the behavior, general factors that might affect the behavior (medications, staffing patterns, daily schedule), antecedents and outcomes of the behavior, functional behavior repertoires, communication skills, potential reinforcers, and treatment history.

A form of the Functional Assessment Interview for students who can act as their own informants is also available (Kern, Dunlap, Clarke, & Childs, 1995; O'Neill et al., 1997). Questions include the behavior or behaviors that cause trouble for the students at school, a description of their class schedule and its relation to problem behavior, rating of intensity of behaviors on a scale of 1 to 6 across class periods and times of day, aspects of the situation related to the behavior (e.g., difficult, boring, or unclear material; peer teasing; teacher reprimands), other events that might affect the behavior (e.g., lack of sleep, conflicts), consequences (what occurs when the individuals engage in the behavior), possible behavior alternatives, and potential strategies for a support plan.

Two other questionnaires are the Behavioral Diagnosis and Treatment Information Form (Bailey & Pyles, 1989) and the Stimulus Control Checklist (Rolider & Van Houten, 1993), which also address questions about the conditions under which the behavior occurs or does not occur and how often. In addition, they include questions about physiological factors that might affect the behavior.

Behavior Rating Scales

Behavior rating scales designed for functional assessment ask informants to estimate the extent to which behavior occurs under specified conditions, using a Likert scale (e.g., *never, seldom, usually, always*). Hypotheses about the function of a behavior are based on the scores associated with each condition. Those conditions assigned the highest cumulative or average rating are hypothesized to be related to the problem behavior. For example, if an informant states that problem behavior always occurs when demands are placed on a child, a negative reinforcement hypothesis might be made. Table 27.2 summarizes features of several behavior rating scales.

Considerations for Indirect FBA

Some indirect assessment methods can provide a useful source of information in guiding subsequent, more objective assessments, and contribute to the development of hypotheses about variables that might occasion or maintain the behaviors of concern. Because indirect forms of FBA do not require direct observation of problem behavior, most behavior analytic practitioners view them as convenient and consider them useful in conducting an FBA. However, we caution readers that closed-ended indirect assessment instruments such as the MAS and QABF have been repeatedly demonstrated to be unreliable and therefore of questionable validity in identifying function (see Hanley, 2012).

CONDUCTING A FUNCTIONAL BEHAVIOR ASSESSMENT

FBA can best be viewed as a four-step process:

1. Gather information with indirect and descriptive assessment.
2. Interpret information from indirect and descriptive assessment and formulate hypotheses about the purpose of problem behavior.
3. Test hypotheses using functional analysis.
4. Develop intervention options based on the function of problem behavior.

Gathering Information

It is often helpful to begin the FBA process by conducting Functional Assessment Interviews with the person's teacher, parent, caregiver, and/or others who work closely with the person. The interview can be helpful in preparing the evaluator to conduct direct observations by identifying and defining the target problem behaviors, identifying and defining potential

TABLE 27.2 Behavior Rating Scales Used to Assess Possible Functions of Problem Behavior

Behavior rating scale	Functions assessed	Format and number of items	Example item and possible function
Motivation Assessment Scale (MAS) (Durand & Crimmins, 1992)	Sensory reinforcement, escape, attention, and tangible reinforcement	16 questions (4 for each of 4 functions), 7-point scale from *always* to *never*	Does the behavior seem to occur in response to your talking to other persons in the room? (attention)
Motivation Analysis Rating Scale (MARS) (Wieseler, Hanson, Chamberlain, & Thompson, 1985)	Sensory reinforcement, escape, and attention	6 statements (2 for each of 3 functions), 4-point scale from *always* to *never*	The behavior stops occurring shortly after you stop working or making demands on this person. (escape)
Problem Behavior Questionnaire (PBQ) (Lewis, Scott, & Sugai, 1994)	Peer attention, teacher attention, escape/avoid peer attention, escape/avoid teacher attention, and assessment of setting events	Questions, 7-point range	When the problem behavior occurs, do peers verbally respond to or laugh at the student? (peer attention)
Functional Analysis Screening Tool (FAST) (Iwata & DeLeon, 1996)	Social reinforcement (attention, preferred items), social reinforcement (escape), automatic reinforcement by sensory stimulation, automatic reinforcement by pain attenuation	Yes or no as to whether statements are descriptive	When the behavior occurs, do you usually try to calm the person down or distract the person with preferred activities (leisure items, snacks, etc.)? (social reinforcement, attention, preferred items)
Questions About Behavioral Function (QABF) (Paclawskyj, Matson, Rush, Smalls, & Vollmer, 2000)	Attention, escape, nonsocial, physical, tangible	Statements, 4-point range	Participant engages in the behavior to try to get a reaction from you. (attention)

antecedents and consequences that may be observed, and gaining an overall picture of the problem behavior as well as the strengths of the person. The interview can also help determine if other assessments are warranted before a more extensive FBA is conducted. For example, if the interview reveals that the person has chronic ear infections that are currently untreated, a medical evaluation should be conducted before further behavioral assessment takes place.

In many cases, conducting an interview with the person who has problem behavior can be helpful if he has the language skills to understand and respond to interview questions. Sometimes the person has useful insights regarding why he displays problem behavior in specific contexts.

At this point, conducting direct observations of the problem behavior within the natural routine is useful. Such observations help confirm or disconfirm the information obtained through the interviews. If it is not clear when the problem behavior occurs most often, a scatterplot analysis may be useful to determine when further behavioral observations should be conducted. When the problematic time periods have been determined, the behavior analyst can conduct ABC assessments. Information obtained from the interviews is helpful in guiding the ABC assessment because the evaluator should already have clear definitions of the target behavior(s), antecedents, and consequences. However, the behavior analyst must also watch for additional, unexpected antecedents and consequences that might present themselves in the natural environment. Teachers or caregivers sometimes overlook or are unaware of specific stimuli triggering or following problem behavior.

Interpreting Information and Formulating Hypotheses

Results from indirect assessments should be analyzed for patterns of behavior and environmental events so that hypotheses regarding the function of the problem behavior can be made. If problem behavior occurs most frequently when low levels of attention are available and problem behavior frequently produces attention, a hypothesis that attention maintains the problem behavior is appropriate. If the problem behavior occurs most frequently in high-demand situations and often produces a reprieve from the task (e.g., through time-out, suspension, or another form of task delay), then a hypothesis that escape maintains the problem behavior is appropriate. If problem behavior occurs in an unpredictable pattern or at high rates across the school day, a hypothesis that the behavior is maintained by automatic reinforcement may be appropriate. In reviewing assessment results and considering possible hypotheses, behavior analysts should remember that behaviors may serve multiple functions and different topographies of problem behavior may serve different functions.

Hypothesis statements should be written in ABC format. Specifically, the hypothesis statement should state the antecedent(s) hypothesized to trigger the problem behavior, the topography of problem behavior, and the maintaining consequence. An example is presented below:

Hypothesized Function	Antecedent	Behavior	Consequence
Escape from hand washing and/or lunch.	When Flo is prompted to wash her hands in preparation for lunch, . . .	she screams and falls to the ground, which is followed by . . .	termination of hand washing and lunch by being sent to time-out.

Writing hypothesis statements in this manner is useful because it requires the behavior analyst to focus on potential avenues for intervention: modifying the antecedent and/or modifying the reinforcement contingencies (which may involve teaching a new behavior and altering what behaviors are reinforced or placed on extinction).

Testing Hypotheses

After hypotheses have been developed, a functional analysis can be conducted to test them. The functional analysis should always contain a control condition that promotes the lowest frequency of problem behavior. Often this is the play condition, which consists of (a) continuous availability of preferred toys and/or activities, (b) no demands, and (c) continuously available attention. Then, conditions are selected to test specific hypotheses. For example, if the primary hypothesis is that the problem behavior is maintained by escape, then a contingent escape condition should be implemented. Other test conditions may not need to be implemented.

Being selective about the test conditions implemented will help keep the functional analysis as brief as possible; however, no conclusions can be made regarding additional functions of problem behavior if additional test conditions are not implemented. For example, if play and contingent escape are the only conditions tested, and problem behavior occurs most frequently in the contingent escape condition and seldom or never occurs in the play condition, the conclusion that escape maintains the problem behavior is supported. However, because a contingent attention condition was not implemented, one could not rule out the possibility that attention also maintains the problem behavior.

Developing Interventions

When an FBA has been completed, an intervention that matches the function of problem behavior can be developed. Interventions can take many forms. Although FBA does not identify which interventions will be effective in treating problem behavior, it does identify antecedents that may trigger problem behavior, potential behavioral deficits that should be remedied, and reinforcement contingencies that can be altered, as described earlier in this chapter. In addition, FBA *does* identify powerful reinforcers that can be used as part of the intervention package. The intervention should be **functionally equivalent** to problem behavior. That is, if problem behavior serves an escape function, then the intervention should provide escape (e.g., in the form of breaks from task demands) for a more appropriate response or involve altering task demands in a fashion that makes escape less reinforcing.

One effective way to design interventions is to review confirmed hypotheses to determine how the ABC contingency can be altered to promote more positive behavior. For example, consider the hypothesis developed for Flo, and assume the functional analysis revealed a pattern similar to that seen in the upper right-hand corner of Figure 27.2:

Hypothesized Function	Antecedent	Behavior	Consequence
Escape from hand washing and/or lunch.	When Flo is prompted to wash her hands in preparation for lunch, . . .	she screams and falls to the floor, which is followed by . . .	termination of hand washing and lunch by being sent to time-out.

The antecedent could be altered by changing the time of day when Flo is asked to wash her hands (so that it does not precede lunch, thereby decreasing the motivation for escape-motivated tantrums).

Hypothesized Function	Antecedent	Behavior	Consequence
Escape from hand washing and/or lunch.	Flo is prompted to wash her hands before recess.	N/A (Problem behavior is avoided.)	N/A (The consequence is irrelevant because problem behavior did not occur.)

The behavior could be altered by teaching Flo a new behavior (e.g., signing "break") that results in the same outcome (escape from lunch).

Hypothesized Function	Antecedent	Behavior	Consequence
Escape from hand washing and/or lunch.	When Flo is prompted to wash her hands in preparation for lunch, . . .	Flo is prompted to sign "break," which is followed by . . .	termination of hand washing and lunch.

Or, the consequences could be altered. For example, the reinforcer for problem behavior could be withheld so that problem behavior is extinguished.

Hypothesized Function	Antecedent	Behavior	Consequence
Escape from hand washing and/or lunch.	When Flo is prompted to wash her hands in preparation for lunch, . . .	she screams and tantrums, which is followed by . . .	continued presentation of hand washing and lunch activities.

An intervention can also consist of several different components. For example, Flo could be taught a replacement behavior (signing "break"), which results in breaks from lunch, while tantrums are simultaneously placed on escape extinction.

FBA can also help identify interventions that are likely to be ineffective or that may worsen the problem behavior. Interventions involving time-out, in-school or out-of-school suspension, or planned ignoring are contraindicated for problem behaviors maintained by escape. Interventions involving reprimands, discussion, or counseling are contraindicated for problem behaviors maintained by attention.

We offer a final word about intervention: When an intervention has been developed, FBA is not "done." Assessment is ongoing once intervention is implemented, as it is important for the continued monitoring of intervention effectiveness. The functions of behavior are not static. Rather, they are dynamic and change over time. Intervention may lose its effectiveness because the function of problem behavior may change (Lerman, Iwata, Smith, Zarcone, & Vollmer, 1994). In such cases, additional functional analyses may be needed to revise the intervention.

CASE EXAMPLES ILLUSTRATING THE FBA PROCESS

FBA is a highly idiosyncratic process. It is unusual for any two FBAs to be exactly the same because each person presents with a unique set of skills and behaviors, as well as a unique history of reinforcement. FBA requires a thorough understanding of behavioral principles to parcel out the relevant information from interviews and ABC assessments, to form relevant hypotheses, and to test those hypotheses. Beyond these skills, a solid understanding of behavioral interventions (e.g., differential reinforcement procedures, schedules of reinforcement, and tactics for promoting maintenance and generalization) is needed to match effective treatments to the function of challenging behavior. Furthermore, it is important for the practitioner to keep abreast of the current literature in functional analysis and treatment of problem behavior, because the evidence base is constantly evolving. This can seem like a daunting process. In an attempt to demonstrate the application of FBA across the idiosyncratic differences in people, we present an array of case examples that illustrate its utility in developing interventions that render the problem behavior ineffective, irrelevant, or inefficient.

Marie—Brief Functional Analysis

Gathering Information, Interpreting Information, and Forming Hypotheses

Marie was an 8-year-old girl with autism who had been referred to a behavior analyst for an evaluation of severe problem behavior, specifically aggression toward her family members. Her insurance covered 3 hours of "ABA therapy" per week, which she was receiving from a community mental health agency social worker. Her mother and sister attended every session.

The only behavior analyst with the skills necessary to consult on this case was 150 miles away at a local university, so the behavior analyst helped the agency staff conduct the functional analysis via teleconsultation.

The behavior analyst began by conducting a Functional Assessment Interview with Marie's mother, who reported that aggression occurred in a variety of settings, toward children and adults, and typically occurred when Marie did not get what she wanted, usually a preferred toy. Aggression also took place during her therapy sessions when the therapist was talking to Marie's mother, updating her on the session. However, aggression toward her on-site therapist had recently decreased. The behavior analyst then collected some ABC data during a play session with Marie's mother and younger sister, as well as during her ABA therapy session at the clinic. The ABC data showed a pattern of aggression when Marie's preferred items were removed. In addition, the ABC data indicated that the consequence that most often followed aggression was Marie's mother or sister relinquishing the item back to Marie. Based on the descriptive data, it was hypothesized that Marie's aggression could be maintained by access to tangibles.

Testing Hypotheses

Due to the limited number of hours per week Marie received services, a brief functional analysis was conducted to determine the function of Marie's aggression. The consulting behavior analyst trained Marie's therapist to conduct the assessment via teleconsultation, and was available during the assessment to provide coaching and feedback. Four 5-min functional analysis conditions (free play, attention, demand, and tangible) were conducted, and aggression occurred only in the tangible condition. During each session, the behavior analyst scored the occurrence of problem behavior using a 10-sec partial interval recording procedure. During the contingency reversal, the number of requests was also recorded. The functional analysis results are shown in Figure 27.5. After the first four conditions, the behavior analyst coached the on-site therapist on how to implement functional communication training (FCT) to teach Marie to ask for the tangible item. This was selected as a replacement behavior because it matched the function of the problem behavior and could efficiently compete for the reinforcer. During three 5-min training sessions, requests increased and aggression dropped to zero. The behavior analyst instructed the on-site therapist to conduct a contingency reversal by implementing a second tangible functional analysis condition. During this reversal, aggression increased to 50% of intervals, and requesting decreased from 11 to 2 occurrences. Upon reinstatement of FCT, aggression returned to zero and requesting returned to the previous treatment level. All functional analysis and treatment conditions were completed in a 60-min appointment with Marie's therapist, mother, and sister present.

Developing an Intervention

During the brief FA, functional communication training (FCT; teaching appropriate requesting) was shown to be an effective treatment for decreasing aggression. It matched the function and was more efficient than the problem behavior in obtaining access

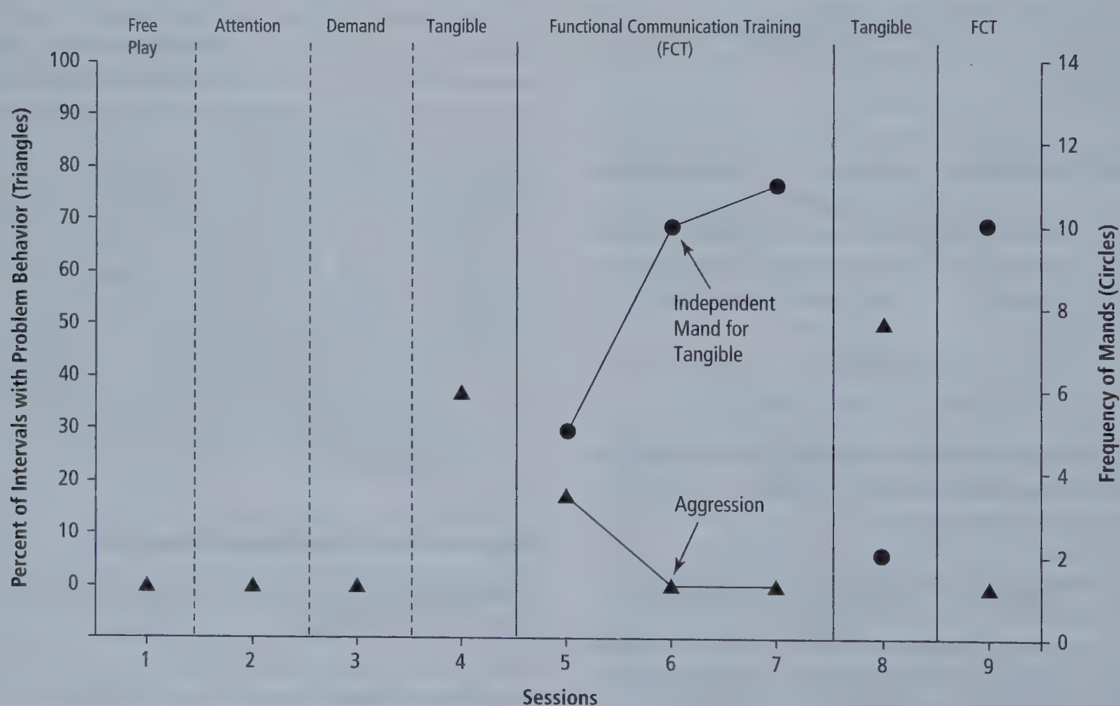


Figure 27.5 Results from the brief functional analysis for Marie. The graph depicts the percentage of 10-sec intervals in which problem behavior occurred (left *y*-axis) and the number of independent mands (right *y*-axis) in each condition.

Marie's case example contributed by Rebecca Eldrige and Nathan VanderWeele.

to the tangible item. During a follow-up meeting, the behavior analyst, the on-site therapist, and Marie's mother met to develop the treatment plan. Marie's mother was concerned about the times that she would need to say "no" to Marie when Marie was asking for preferred items and was worried that aggression would result. The behavior analyst worked with the team to develop a treatment plan that included initially delivering the item each time Marie requested it until aggression had decreased to zero for three consecutive 30-min sessions. When Marie was regularly making requests without aggression, a program to teach waiting for the preferred item was introduced. To gain access to the preferred item, Marie had to wait (without aggression) for 1 to 60 seconds after requesting it. Once waiting was mastered, Marie was taught to complete a task while while waiting. She could then earn the preferred item only after the task was complete. The behavior analyst and on-site therapist then taught Marie's mother to implement the procedure, which she did successfully.

Carson—Trial-based Functional Analysis

Gathering Information, Interpreting Information, and Forming Hypotheses

Carson, a 10-year-old boy with a chromosomal deletion disorder, was receiving special education services in the general education classroom. He displayed several problem behaviors—most notably, talking to peers during group instruction, yelling out answers, and running around the room. These behaviors occurred only in the general education classroom. They had increased in intensity, resulting in more time spent in the resource room, away from the general education setting. Consequently, Carson was missing valuable instruction time and falling behind in the curriculum.

The behavior analyst assigned to Carson's case was asked to complete an FBA in order to design an intervention plan that would increase the time he spent in the general education classroom and decrease his disruptive behavior. After obtaining consent from Carson's parents to conduct an FBA, the behavior analyst conducted a Functional Assessment Interview with Carson's teachers and paraprofessionals. From the interview, it appeared that Carson's problem behaviors occurred most often during group lessons on the carpet and individual work at his desk. Carson's teachers and paraprofessionals could not identify a clear pattern of consequences for his behavior because they had tried so many different strategies without success. Following the interview, the behavior analyst used a structured data form to begin collecting ABC data during those times reported as most problematic by the teachers. The ABC assessment showed that escape from demands or access to one-on-one attention could be maintaining Carson's disruptive behavior.

Testing Hypotheses

At this point, the behavior analyst needed to test her hypotheses, but she had several reservations about conducting a traditional FA in the classroom setting. She was concerned about the time that a traditional FA would take as well the teacher's ability to implement it with a high degree of fidelity while also teaching. For these reasons, the behavior analyst decided to conduct a trial-based functional analysis (TBFA) on disruptive behavior, with the teacher implementing the conditions. Four conditions were conducted: attention test, attention control, escape test, and escape control. During the attention test condition, teacher attention was diverted unless Carson engaged in disruption. If he engaged in disruption, the teacher provided attention for 10 seconds and ended the trial. During the attention control condition, attention was delivered every 5 seconds for 60 seconds, with no programmed consequences for disruptive behavior. During the escape test condition, the teacher worked one-on-one with Carson, repeatedly prompting him to complete his work. If he engaged in disruption, the teacher walked away and removed task demands for 10 seconds. During the escape control condition, the teacher provided a 1-min break with no attention or task demands and no programmed consequences for disruption. Each condition was implemented 8 times (trials) for 1 min during regular instruction throughout the school day. The percentage of trials with disruption in each condition was then calculated and graphed. The results (see Figure 27.6) showed that disruption occurred most often in the attention

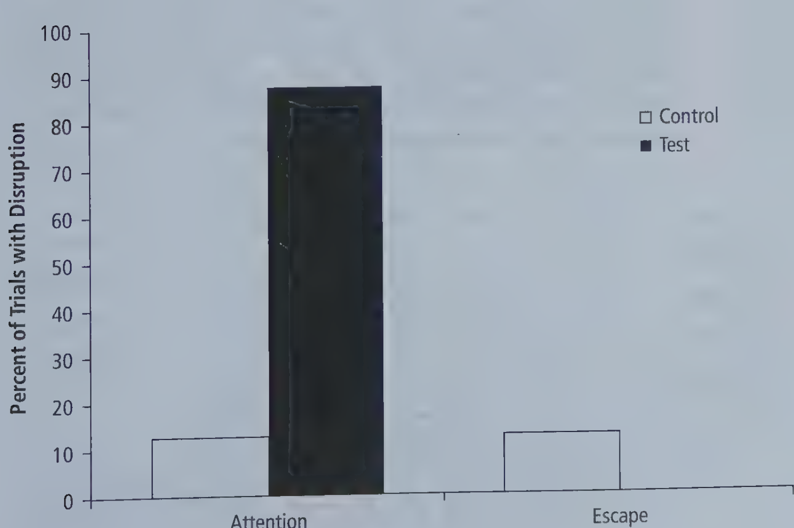


Figure 27.6 Results from the trial-based functional analysis for Carson. The graph shows the percentage of trials, for both the control and the test conditions, in which disruptive behavior occurred.

Carson's case example contributed by Rebecca Eldrige and Nathan VanderWeele.

test condition, and thus supported the hypothesis that attention was maintaining disruptive behavior. In addition, the escape test condition evoked no problem behavior, which suggested that escape from demands was not a reinforcing variable for disruptive behavior.

Developing an Intervention

Based on the results of the trial-based functional analysis, a multicomponent treatment plan was developed that included discontinuing attention when Carson engaged in disruptive behavior, teaching him to raise his hand to gain attention, and providing contingent breaks with teacher attention for sitting quietly at his desk for increasing amounts of time. Thus, problem behavior was placed on extinction, and more appropriate behaviors (hand raising and sitting quietly engaged in his work) produced teacher attention.

Elija—Latency-based Functional Analysis

Gathering Information, Interpreting Information, and Forming Hypotheses

Elija was a 9-year-old boy diagnosed with autism spectrum disorder. He engaged in elopement and noncompliance and was referred to a behavior analyst for a functional analysis.

The behavior analyst began with a Functional Assessment Interview (O'Neill et al., 1997) with Elija's mother. Elija's mother reported that he typically engaged in elopement in the community as well as in the home; he would run out of the home or store and into the street when he wanted something he couldn't have (while in the community) or when he was denied access to his toys (at home). Consequently, his mother did not take him out much, which limited the number of activities Elija could engage in throughout the day. His mother kept all the doors locked in the house.

Next, the behavior analyst conducted some direct observation of Elija's behavior at home using remote technology. Elija's mother was observed working with him under a variety of naturally occurring conditions. The behavior analyst took data on elopement attempts as well as the immediate antecedents and consequences. Attempts to elope were blocked to keep Elija safe. During the ABC assessment, the behavior analyst observed elopement attempts when Elija was asked to complete academic tasks. Additionally, when he eloped, Elija's mother provided extensive attention in an attempt to get him to sit back down at the table and do his work. Based on the ABC observations, the behavior analyst hypothesized that Elija's elopement was maintained by escape from task demands as well as access to tangible items, and attention.

Testing Hypotheses

The functional analysis was conducted at the center where Elija received services. The behavior analyst decided to conduct a latency-based functional analysis. She chose this assessment given the dangerous nature of the problem behavior. That is, she wanted to minimize the number of occurrences of elopement, and staff had to intervene when elopement did occur. The behavior analyst immediately reinforced elopement, which prevented Elija from eloping too far from the setting. Elopement was defined as any instance of Elija moving more than 3 feet away from the instructional setting and/or from an adult without permission. In addition to conducting the latency FA, the behavior analyst also ensured that additional therapists were assigned to all exits in the building to prevent Elija from going out of the building. The FA consisted of the following conditions: tangible, play (control), attention, and demand. The behavior analyst started a stopwatch when the condition began and let it run until elopement occurred. When elopement occurred, the behavior analyst stopped the stopwatch and recorded the time that had elapsed. This was the latency to elopement. The results of the FA are shown in Figure 27.7. The graph depicts that elopement was maintained by escape from demands and access to attention, because Elija's latency to elopement was shortest during the escape and attention conditions.

Developing an Intervention

It can be challenging when elopement or any other dangerous problem behavior serves an attention function, because the problem behavior cannot be placed on extinction (i.e., ignored) safely. The problem behavior will likely always result in at least some attention. Such was the case for Elija. As a result, the behavior analyst designed an intervention based on concurrently available reinforcement contingencies. When Elija was presented with a task demand, he was asked to make a choice from among three concurrently available options, which were signaled by three different colors (green, yellow, and red). Green was associated with Elija completing the assigned task demand. If he appropriately completed the task, he was given a high-quality break for 60 seconds. During his break, he received attention from his mother, as well as preferred toys. If he asked for a break from the

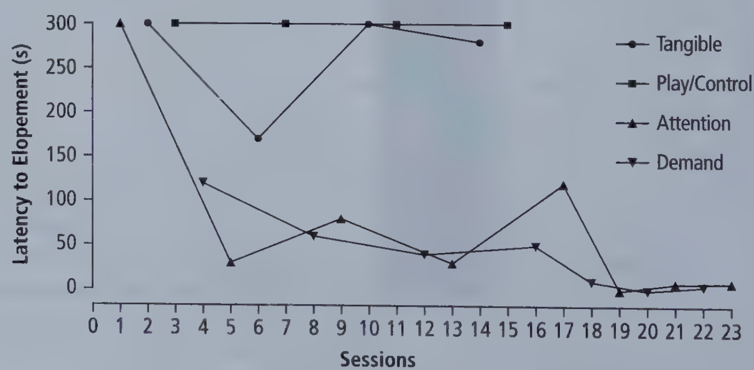


Figure 27.7 The latency to elopement for each condition in the FA for Elija. Note that elopement never occurred in the Play/Control condition; thus, the data points are plotted at 300 s, which represents the entire session duration (5 min). The latency was shortest during the attention and demand conditions, indicating escape and attention functions.

Elija's case example contributed by Denice Rios and Nicole Hollins.

task (the yellow option), he was given a task break. However, the break lasted only 30 seconds, he had access to less preferred toys, and his mother stayed with him for only a few seconds. If he engaged in elopement (the red option), he escaped the task for 10 seconds, no toys were available, and his mother neutrally blocked his egress from the room. Over time, the duration of the task demand to earn the high-quality break increased until Elija reliably engaged in the task demand rather than eloping.

Will—Evaluating Precursor Behavior

Gathering Information, Interpreting Information, and Forming Hypotheses

Will was 14 years old with a primary diagnosis of autism. He was quite large for his age and displayed high levels of physical aggression, verbal outbursts, and urinating on people. As a result of the severity and frequency of his problem behaviors, his size, a lack of informed treatment, and limited resources, Will was sent to a high-security juvenile detention center, where he was exposed to individuals without diagnosed disabilities who had committed violent, serious, and/or gang-related crimes. The county serving Will recognized that the juvenile detention center was not an appropriate placement for him but had no other placement options. As a last resort, the service coordinator on the case sought a behavior analytic consultation.

Upon getting involved with the case, the behavior analyst realized that everyone working with Will was skeptical that anything would help. The behavior analyst conducted an FBA interview with individuals who had worked with Will in the detention center. Because the center was a lock-down facility, the staff could not give much information about how Will performed in a variety of conditions. The interview yielded little useful information. Also, observations of natural routines were limited due to the nature of the placement. As a result, no formal descriptive assessments were completed. The behavior analyst proposed a functional analysis, but one of Will's caseworkers was concerned that once Will got upset, he would engage in aggression for hours. She was also concerned that Will's behaviors were so severe that allowing them to occur even once could be dangerous.

Based on a review of Will's records and the limited information available, it was hypothesized that Will's aggressive behaviors were maintained by escape from demands or by attention. In addition, it was determined that verbal outbursts describing the aggressive behavior he was about to engage in (e.g., "I'm going to piss on you!") reliably preceded Will's aggression. As a result of all these factors, it was decided that the functional analysis would be conducted on Will's verbal outbursts rather than on his aggression.

Testing Hypotheses

The functional analysis consisted of four conditions: tangible, free play, escape, and attention. In each of these conditions, occurrences of verbal outbursts were recorded using a 10-sec partial interval measurement system and plotted as the percentage of intervals with problem behavior (see Figure 27.8). Will displayed high levels of verbal outbursts during the tangible condition. The behavior analyst noted that Will's verbal outbursts could be interrupted if the therapist provided attention or task demands (i.e., the attention and demand conditions that followed the tangible condition) that redirected Will's behavior to some form of independent work.

Developing an Intervention

Based on the results of the functional analysis, the behavior analyst recommended that Will receive training in effectively and appropriately asking for tangible items. In addition, it was recommended that the training focus on teaching Will to tolerate delays in getting access to those tangible items. In his typical environment (e.g., school and home setting), gaining immediate access to tangible items was not always going to be feasible, so tolerating delays would be imperative for Will's success. Thus, a structured treatment that involved providing access to tangible items following appropriate requests, but not verbal outbursts, was recommended. After Will reliably requested items in an appropriate way, the behavior analyst slowly built in delays to the preferred item and asked Will to work on an independent task while waiting. Following appropriate waiting, Will received access to the preferred item. After observing the process, Will's caseworker—albeit initially resistant—considered that Will could live in a less restrictive environment.

Chris—Evaluating ABC Data

Gathering Information, Interpreting Information, and Forming Hypotheses

Chris was a 37-year-old man diagnosed with intellectual disabilities who lived in a group home. He had a limited vocal verbal repertoire and he used a wheelchair for mobility. Chris displayed

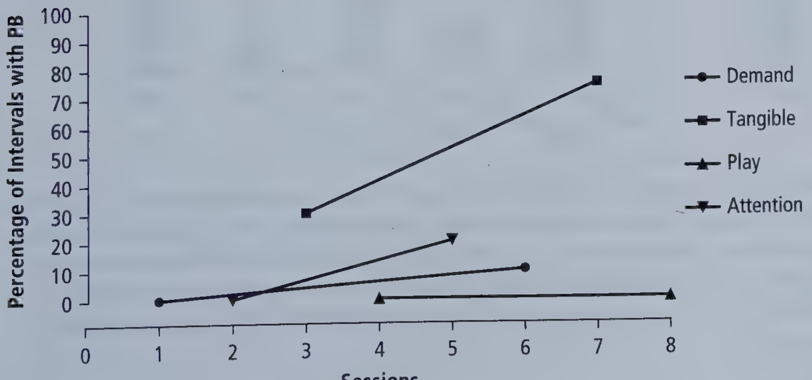


Figure 27.8 Results of Will's functional analysis of precursor behavior (verbal outbursts).

Will's case example contributed by Cody Morris, Denice Rios, and Lloyd Peterson.

frequent screaming toward caregivers and peers across many locations, including his home, various day programs, and outings. Interviews with staff members indicated that the screaming occurred throughout the day, but most frequently at his day programs. The staff hypothesized that the screaming occurred when Chris observed other roommates getting attention from staff. The behavior analyst also noticed that when staff members moved Chris's wheelchair, they often did so by walking up behind him and moving his wheelchair without saying anything to Chris about what was going to happen. This seemed to startle Chris, and the behavior analyst wondered if his screaming might also be maintained by avoidance of being removed from a specific area in the day program.

Structured ABC observations (similar to that in Figure 27.3) and narrative recordings (providing descriptors of each event) were conducted across various times in Chris's home and day programs. Both inappropriate vocalizations (screaming) and appropriate vocalizations were tracked. After data were collected, each antecedent and consequence entry was coded into categories, such as attention, presentation or removal of a demand, being relocated, and alone. Each occurrence of appropriate and inappropriate vocalization was then analyzed to determine what category of antecedent immediately preceded it and what category of consequence followed it. The percentage of both appropriate and inappropriate vocalizations preceded or followed by each category of variables was then graphed and analyzed using methods described in Chapter 6. The results are shown in Figure 27.9.

Based on the ABC observations and graphed ABC data, it was hypothesized that Chris's screaming served an attention function. The graphed data clearly showed that the most common antecedent variable that preceded both appropriate and inappropriate vocalizations was Chris being alone. However, Chris rarely received attention following appropriate attempts to interact with staff. In contrast, he reliably received staff attention following screaming. In fact, Chris was about twice as likely to receive attention for inappropriate vocalizations as for appropriate ones. Moreover, the narrative aspect of the ABC recording (see Table 27.3) indicated that staff members reliably ignored Chris while he scrolled through a wide variety of appropriate vocalizations (e.g., "hi," "how are you") that escalated to grunting and eventually screaming. Typically, it wasn't until Chris began screaming that he received high-quality attention from staff.

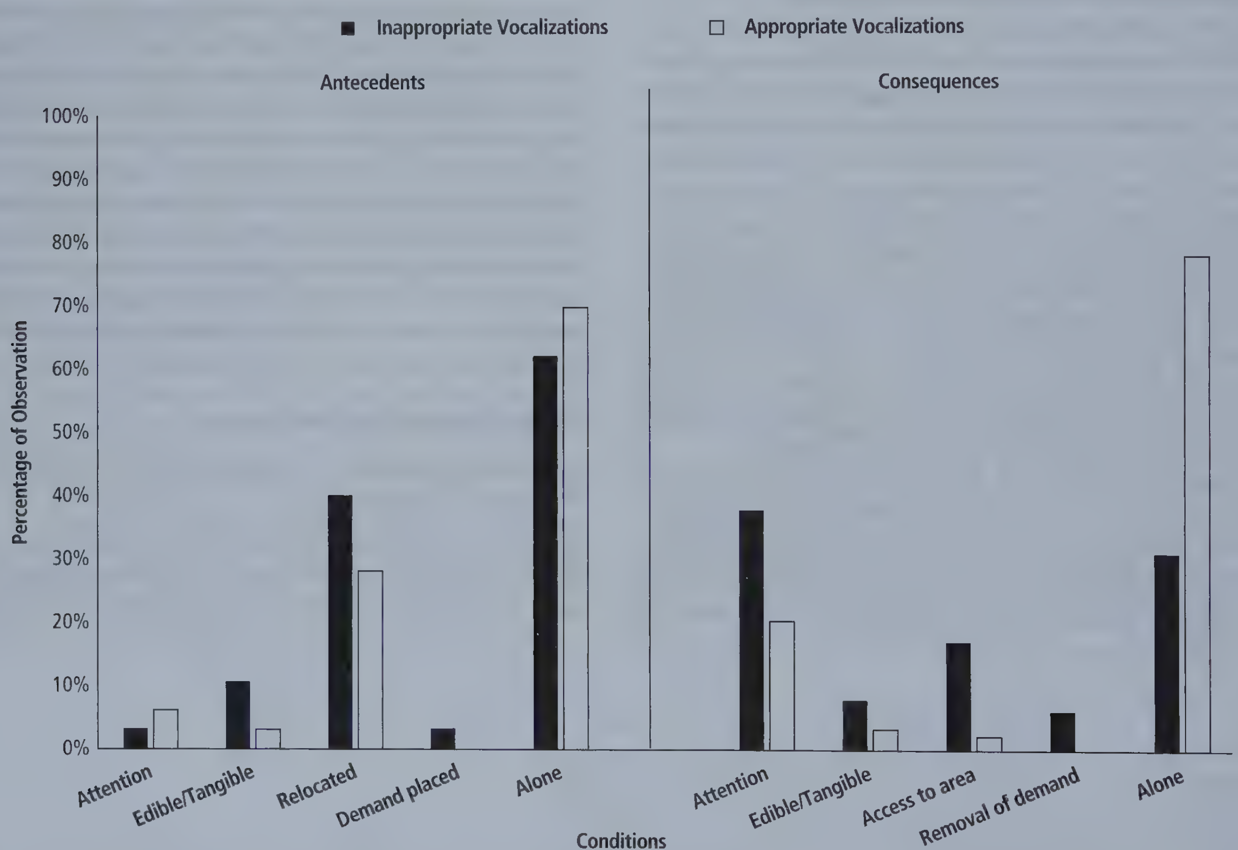


Figure 27.9 Results of Chris's ABC assessment. The left-hand panel shows the percentage of appropriate and inappropriate vocalizations that were preceded by the antecedent variables of one-on-one attention, access to edibles/tangibles, having someone move his wheelchair (relocated), having a demand placed on him, or being left alone. The right-hand panel shows the percentage of appropriate and inappropriate vocalizations that were followed by the consequence variables of one-on-one attention, access to edibles/tangibles, having access to a new area, having a demand removed, or being left alone.

Case example contributed by Cody Morris and Becky Kolb.

TABLE 27.3 Results of ABC Narrative Assessment for Chris's Appropriate and Inappropriate Behavior

Antecedent	Behavior	Consequence
Chris alone	Said "hi you"	Ignored by staff
Chris alone	Said "a ba" while pointing to a book	Ignored by staff
Chris alone	Screamed	Approached and told "Please stop screaming" and asked "What's wrong, Chris?"
Chris transitioned to a new activity (sorting)	Screamed	Staff asked Chris why he was upset—"Chris, what do you need?"
Chris transitioned to a different room (because of an event occurring)	Said "are you my boy" to staff while walking	Ignored by staff

Developing an Intervention

Because the ABC analysis identified a very apparent issue with staff interaction with Chris, it was not clear that a functional analysis was necessary. The behavior analyst wanted to first try having staff attend to appropriate behavior that Chris was already displaying. The intervention consisted of reversing staff's differential reinforcement of the inappropriate vocalizations. This was done by having staff reinforce the appropriate vocalizations (i.e., responding to the appropriate bids for attention) while placing screaming on extinction. In addition, because Chris engaged in high levels of appropriate vocalizations already, he was taught to tolerate a delay to reinforcement. Staff members were asked to acknowledge his requests and to provide reinforcement (attention) immediately, whenever possible. However, staff members were also taught to work directly with Chris to teach him to "wait" when necessary by slowly increasing delays to attention following their acknowledgement of his request. Finally, because Chris's vocalizations were somewhat difficult to understand and limited in variety, interventions to expand and improve the quality of Chris's mand repertoire were introduced.

Brian—Multiple Functions of Problem Behavior

Gathering Information

Brian was 13 years old and diagnosed with pervasive developmental delay, oppositional defiant disorder, and attention-deficit hyperactivity disorder. He had moderate delays in cognitive and adaptive skills. Brian displayed several problem behaviors, including aggression, property destruction, and tantrums. Brian's aggression had resulted in several of his teachers having bruises, and his property destruction and tantrums frequently disrupted the daily activities of the classroom.

A Functional Assessment Interview (O'Neill et al., 1997) was conducted with Brian's teacher, Ms. Baker, who reported that Brian's problem behavior occurred most frequently when he was asked to perform a task that required any kind of physical labor (e.g., shredding papers) and occurred least during leisure activities. However, Ms. Baker reported that Brian often engaged in problem behavior when he was asked to leave a preferred activity. She noted that Brian used complex speech (sentences), although he often used verbal threats, curse words, and/or aggression, property destruction, and tantrums to communicate his wants and needs.

Because Brian had a spoken repertoire, a Student-Assisted Functional Assessment Interview (Kern et al., 1995) was also conducted. In this interview, Brian reported that he found his math work too difficult but that writing and using a calculator were too easy. He reported that he sometimes received help from his teachers when he asked for it, that sometimes teachers and staff noticed when he was doing a good job, and that he sometimes received rewards for doing good work. Brian indicated that his work periods were always too long, especially those that consisted of shredding papers. Brian reported that he had the fewest problems in school when he was allowed to answer the phone (his classroom job), when he was completing math problems, and when he was playing with his Gameboy. He stated that he had the most problems at school when he was outside playing with the other students because they often teased him, called him names, and cursed at him.

An ABC assessment was conducted on two separate occasions. The results of the ABC assessment are shown in Table 27.4.

Interpreting Information and Formulating Hypotheses

Based on the interviews and ABC assessments, the function of Brian's problem behavior was unclear. It was hypothesized that some of Brian's problem behaviors were maintained by access to adult attention and preferred items. This hypothesis was a result of the ABC assessment, which indicated that many of Brian's problem behaviors occurred when adult attention was low or when access to preferred items was restricted. Brian's problem behavior often resulted in access to adult attention or preferred activities. It was also hypothesized that Brian's problem behavior was maintained by escape because his teacher reported that Brian frequently engaged in problem behavior in the presence of task demands and because Brian reported that some of his work was too hard and work periods lasted too long. Thus, a functional analysis was conducted to test these hypotheses.

Testing Hypotheses

The functional analysis consisted of the same conditions as described previously, with two exceptions. First, an alone condition was not conducted because there was no reason to believe that Brian's problem behavior served an automatic function. Second, a contingent tangible condition was added because there was reason to believe Brian engaged in problem behavior to gain access to preferred tangibles and activities.

TABLE 27.4 Results of ABC Assessments for Brian's Aggression, Property Destruction, and Tantrums

Antecedent	Behavior	Consequence
Adult attention diverted to another student; denied access to Nintendo by teacher (i.e., told no when he asked if he could play it)	Yelled at teacher, "That's not fair! Why do you hate me?!"	Told to "calm down"
Teacher attending to another student	Hit sofa, attempted to leave classroom	Given choice of activity and verbal warning to stay in classroom
Teacher attention diverted to another student	Yelled "Stop!" at another student	Reprimand from teacher: "Don't worry, Brian. I will take care of it."
Story time, teacher attending to other students	Laughed loudly	Reprimand from teacher: "Stop it!"
Story time, teacher listening to other students	Interrupted other students while they were talking: "Hey, it's my turn. I know what happens next!"	Reprimand from teacher: "You need to listen."

This condition was just like the play condition (i.e., Brian had access to adult attention and preferred toys at the beginning of the session), except that intermittently throughout the session, he was told it was time to give his toy to the teacher and to play with something else (which was less desirable). If Brian complied with the request to give the toy to the teacher, he was given a less preferred toy. If he engaged in problem behavior, he was allowed to continue playing with his preferred toys for a brief period.

The results of the functional analysis are shown in Figure 27.10. Notice that problem behavior never occurred in the play condition when continuous attention and preferred items were available and no demands were placed on Brian. However, it occurred at high rates in all three of the test conditions (contingent attention, escape, and tangible). These results indicated that Brian's problem behavior was maintained by escape, attention, and access to preferred items.

Developing an Intervention

Based on the results of the functional analysis, a multicomponent intervention was implemented. The intervention components changed at different points in time, depending on the context. For example, when Brian was engaged in a work task, it was recommended that he be given frequent opportunities to request breaks. In addition, the time-out intervention that the teacher had been using was discontinued in the work context. During leisure times, when Brian had previously been expected to play alone, the classroom schedule was rearranged so that Brian could play and interact with peers. Brian was also taught to request toys appropriately while playing with peers. In addition, several interventions aimed at increasing teacher attention for appropriate behavior were implemented. Brian was taught how to request teacher attention appropriately, and teachers began to respond to these requests rather than ignoring them, as they had been doing. In addition, a self-monitoring plan was established, in which Brian was taught to monitor his own behavior and to match his self-recordings to the recordings of his teachers. Accurate self-recording resulted in teacher praise and access to preferred activities with the teacher. Brian's teachers also implemented their own plan to increase attention and praise every 5 minutes as long as Brian was not engaged in problem behavior during independent work.

Lorraine—Multiple Topographies That Serve Multiple Functions

Lorraine was 32 years old and functioned in the moderate range of intellectual disability. She had a diagnosis of Down syndrome and bipolar disorder with psychotic symptoms, for which Zoloft

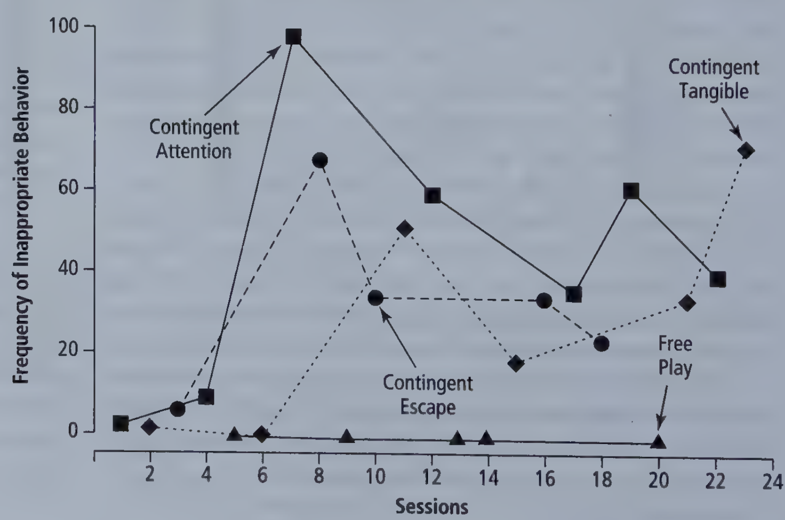


Figure 27.10 Results of Brian's functional analysis. Inappropriate behavior consists of aggression, property destruction, and tantrums.

Based on Brian's functional analysis conducted by Renee Van Norman and Amanda Flaute.

(sertraline) and Risperdal (risperidone) were prescribed. She also took Tegretol (carbamazepine) for seizure control. Her verbal skills were low and her articulation was poor. She communicated through some signs, a simple communication device, gestures, and some words.

Lorraine had resided in a group home for 9 years and attended a sheltered workshop during the day. Lorraine displayed noncompliance, aggression, and self-injurious behavior (SIB) in both settings, but the FBA focused on her problem behavior in the group home, where it was more severe and frequent. Noncompliance consisted of Lorraine putting her head down on the table, pulling away from people, or leaving the room when requests were made of her; aggression consisted of kicking others, throwing objects at others, biting others, and squeezing others' arms very hard; SIB consisted of biting her arm, pulling her hair, or pinching her skin.

Gathering Information

Interviews were conducted with Lorraine, her parents, and workshop and group home staff. Lorraine's parents noted that some of her behavior problems had increased when changes in her medication had been made. Workshop staff noted that Lorraine was more likely to have problem behavior at work if many people were around her. Workshop staff had also noted that noncompliance had increased shortly after a dosage change in medication 2 months previously. The group home staff members noted that they were most concerned about Lorraine's leaving the group home when she was asked to perform daily chores. Lorraine would often leave the group home and not return until the police had picked her up. Many neighbors had complained because Lorraine would sit on their porches for hours until the police came and removed her.

An ABC assessment was conducted at the workshop and group home to determine whether environmental variables differed across the two settings (e.g., the manner in which tasks were presented, the overall level of attention). At the workshop, Lorraine was engaged in a jewelry assembly task (one she reportedly enjoyed), and she worked well for $2\frac{1}{2}$ hours. She appeared to work better when others paid attention to her and often became off task when she was ignored; however, no problem behavior was observed at work. At the group home, aggression was observed when staff ignored Lorraine. No other problem behavior occurred. No demands were placed on Lorraine in the group home during the ABC observation. Group home staff rarely placed any demands on Lorraine in an attempt to avoid her problem behavior.

Interpreting Information and Formulating Hypotheses

Some of Lorraine's problem behaviors seemed to be related to a dosage change in her medication. Because Lorraine's physician judged her medication to be at therapeutic levels, a decision was made to analyze the environmental events related to her problem behavior. Observations of problem behavior during the ABC assessment were limited because workshop staff placed minimal demands on Lorraine to avoid problem behaviors. However, Lorraine's noncompliance reportedly occurred when demands were placed on her. Therefore, it was hypothesized that these problem behaviors were maintained by escape from task demands. Aggression occurred during the ABC assessment when Lorraine was ignored. Although SIB was not observed during the ABC assessment, group home staff reported that Lorraine often engaged in SIB during the same situations that evoked aggression. Therefore, it was hypothesized that both aggression and SIB were maintained by attention.

Testing Hypotheses

The functional analysis consisted of free play, contingent attention, and contingent escape conditions (see Figure 27.11). Because the problem behaviors may have served different functions, each problem behavior was coded and graphed separately. Noncompliance occurred most frequently during the contingent escape condition and rarely occurred during the free play or contingent attention condition. SIB occurred most frequently during the contingent attention condition and rarely occurred during the free play or contingent escape condition. These data suggested that noncompliance served an escape function and SIB served an attention function. As is often the case for low-frequency, high-intensity behaviors, it was difficult to form hypotheses about the function of aggression because the behavior occurred rarely in any of the FBA conditions.

Developing an Intervention

Different interventions were developed for the problem behaviors because results of the FBA suggested that the behaviors served different functions. To address noncompliance, Lorraine was taught to request breaks from difficult tasks. Tasks were broken down into very small steps. Lorraine was presented with only one step of a task at a time. Each time a task request was made, Lorraine was reminded that she could request a break (either by saying "Break, please" or by touching a break card). If she requested a break, the task materials were removed for a brief period. Then they were presented again. Also, if Lorraine engaged in noncompliance, she was *not* allowed to escape the task. Instead, she was prompted through one step of the task, and then another step of the task was presented. Initially, Lorraine was allowed to completely escape the task if she appropriately requested a break each time the task was presented. Over time, however, she was required to complete increasing amounts of work before a break was allowed.

Intervention for aggression consisted of teaching Lorraine appropriate ways to gain attention (e.g., tapping someone on the arm and saying, "Excuse me") and teaching group home staff to regularly attend to Lorraine when she made such requests. In addition, because her articulation was so poor, a picture communication book was created to assist Lorraine in having conversations with others. This communication book could be used to clarify words that staff could not understand. Finally, staff members were encouraged to ignore Lorraine's SIB when it did occur. In the past, staff had approached Lorraine and stopped her from engaging in SIB when it occurred. The functional analysis demonstrated that this intervention may have increased the occurrence of SIB, so this practice was discontinued.

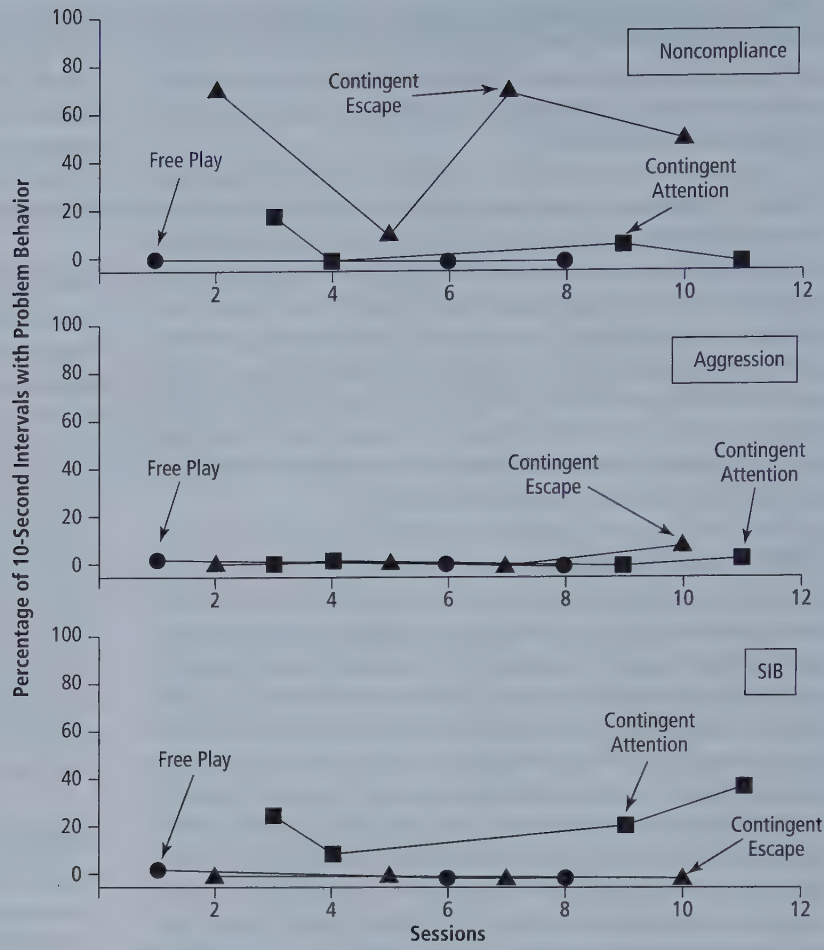


Figure 27.11 Results of Lorraine's functional analysis.

Based on Lorraine's functional analysis conducted by Corrine M. Murphy and Tabitha Kirby.

SUMMARY

Functions of Behavior

1. Many problem behaviors are learned and maintained by positive, negative, and/or automatic reinforcement. In this respect, problem behavior can be said to have a "function" (e.g., to "get" something or to "get out of" something).

2. The *topography*, or form, of a behavior often reveals little useful information about the conditions that account for it. Identifying the *conditions* that account for a behavior (its function) suggests what conditions need to be altered to change the behavior. Assessment of the function of a behavior can therefore yield useful information with respect to intervention strategies that are likely to be effective.

Role of Functional Behavior Assessment in Intervention and Prevention

3. FBA can lead to effective interventions in at least three ways: (a) It can identify antecedent variables that can be altered to prevent problem behavior, (b) it can identify reinforcement contingencies that can be altered so that problem behavior no longer receives reinforcement, and (c) it can help identify reinforcers for alternative replacement behaviors.

4. FBA can decrease reliance on default technologies (increasingly intrusive, coercive, and punishment-based interventions) and contribute to more effective interventions. When FBAs are conducted, reinforcement-based interventions are more likely to be implemented than are interventions that include a punishment component.

Overview of FBA Methods

5. FBA methods can be classified into three types: (a) functional (experimental) analysis, (b) descriptive assessment, and (c) indirect assessment. The methods can be arranged on a continuum with respect to considerations such as ease of use and the type and precision of information they yield.

6. Functional analysis involves systematically manipulating environmental events thought to maintain problem behavior within an experimental design. The primary advantage of functional analysis is its ability to yield a clear demonstration of the variable or variables that relate to the occurrence of a problem behavior. However, this assessment method requires a certain amount of expertise to implement and interpret. Several variations of functional analysis procedures have been developed to adapt to a variety of situations and contexts, including brief functional analysis, conducting functional analyses in natural settings, trial-based functional analysis, synthesized functional analysis, latency-based functional analysis, and analyzing precursor behaviors.

7. Descriptive assessment involves observation of the problem behavior in relation to events that are not arranged in a systematic manner and includes ABC recording (both continuous and narrative) and scatterplot recording. The primary advantages to these assessment methodologies are that they are easier to do than functional analyses and they represent contingencies that occur within the individual's natural routine. Caution must be exercised when interpreting information from descriptive assessments, however, because it can be very difficult to parse the contingencies via them.

8. Indirect functional assessment methods use structured interviews, checklists, rating scales, or questionnaires to obtain information from persons who are familiar with the individual exhibiting the problem behavior (e.g., teachers, parents, caregivers, and/or the individual himself or herself) to identify possible conditions or events in the natural environment that correlate with the problem behavior. Again, these forms of FBA are easy to conduct, but they are limited in their accuracy. As such, they are probably best reserved for hypothesis formulation. Further assessment of these hypotheses is almost always necessary.

Conducting a Functional Behavior Assessment

9. Given the strengths and limitations of the different FBA procedures, FBA can best be viewed as a four-step process:
   - First, gather information via indirect and descriptive assessment.
   - Second, interpret information from indirect and descriptive assessment and formulate hypotheses about the purpose of problem behavior.
   - Third, test hypotheses using functional analysis.
   - Fourth, develop intervention options based on the function of problem behavior.

10. When teaching an alternative behavior as a replacement for problem behavior, the replacement behavior should be functionally equivalent to the problem behavior (i.e., the replacement behavior should produce the same reinforcers that previously maintained the problem behavior).

KEY TERMS

conditional probability
contingency reversal
descriptive functional behavior assessment
functional analysis
functional behavior assessment (FBA)
functionally equivalent
indirect functional assessment

MULTIPLE-CHOICE QUESTIONS

1. Devonia hits her head with a closed fist when her one-on-one teaching assistant leaves her side to interact with another student. Usually, when Devonia does this, her teaching assistant returns to her side, asks her to stop hitting herself, and soothes her. She rarely engages in head hitting when her assistant works directly with her. What is the most likely function of Devonia's problem behavior?
   a. Escape
   b. Attention
   c. Automatic reinforcement
   d. Tangible
   e. Escape and/or tangible
   Hint: (See "Functions of Behavior")

2. Charles spits on his teacher when he prompts him to complete a toothbrushing task. For obvious reasons, this behavior really grosses out the teacher, who refuses to work with him when he behaves like this. When Charles spits on him, the teacher walks away and allows Charles to leave the toothbrushing task and go to the computer to calm down. As long as he is playing on the computer, he rarely spits on his teacher. What is the most likely function of Charles' behavior?
   a. Escape
   b. Automatic reinforcement
   c. Tangible
   d. Escape and/or tangible
   e. Escape and/or attention
   Hint: (See "Functions of Behavior.")

3. On a practical level, FBA is important for prevention of and intervention for problem behavior because:
   a. When the cause-and-effect relation between environmental effects and behavior can be determined, that relation can be altered to improve behavior.
   b. It meets the federal guidelines for best practices in treating problem behavior.
   c. It will tell a teacher exactly what to do for intervention.
   d. None of these
   Hint: (See "Role of Functional Behavior Assessment in Intervention and Prevention")

4. Devonia's one-on-one assistant (see Question 1) decides that he will no longer leave Devonia's side to interact with other children as an intervention for her problem behavior. This is an example of what form of intervention?
   a. Altering antecedent variables
   b. Altering consequent variables
   c. Teaching alternative behaviors
   d. All of these

   Hint: (See "Role of Functional Behavior Assessment in Intervention and Prevention")

5. There are at least three forms of FBA. They are:
   a. Functional analysis, standardized assessment, and descriptive assessment
   b. Indirect assessment, descriptive assessment, and behavioral observation
   c. Functional analysis, descriptive assessment, and indirect assessment
   d. Behavioral observation, standardized assessment, and curriculum-based assessment

   Hint: (See "Overview of FBA Methods")

6. Which methods of FBA allow you to confirm hypotheses regarding the function of problem behavior?
   a. Functional analysis
   b. Descriptive assessment
   c. Indirect assessment
   d. All of these

   Hint: (See "Overview of FBA Methods")

7. Which of the following describes a descriptive functional behavior assessment?
   a. An analog analysis in which consequences representing those in the natural routine are systematically arranged.
   b. Direct observation of behavior made under naturally occurring conditions
   c. Structured checklists that caregivers fill out to identify events that correlate with problem behavior
   d. Behavior rating scales that caregivers fill out to identify events that correlate with problem behavior

   Hint: (See "Descriptive Functional Behavior Assessment")

8. A limitation of descriptive assessment is:
   a. It may be misleading in that it can identify environmental variables that occur in close proximity to the problem behavior but that are not causally related to the problem behavior.
   b. It may not be a very reliable measure of problem behavior and environmental events.
   c. It is extremely difficult and time-consuming to perform.
   d. The first two answer choices are correct.

   Hint: (See "Descriptive Functional Behavior Assessment")

9. Ms. Frieder, who teaches fifth grade at Franklin Elementary School, decides to conduct a descriptive FBA for Amelia. Amelia has been refusing to do her schoolwork lately and has been ripping up her worksheets on a daily basis. Ms. Frieder creates a form on which she will mark specific antecedents and consequences that precede and follow Amelia's work refusals. For antecedents, she will mark one of the following: Math work given, reading work given, spelling work given, error correction given, work that requires writing given. For consequences, she will mark one of the following: verbal reprimand, another worksheet (same worksheet) provided, another worksheet (different worksheet) provided, time out, ignore. She will record these environmental events, whether or not a problem behavior occurs. What kind of descriptive assessment is Ms. Frieder conducting?
   a. Scatterplot recording
   b. ABC continuous recording
   c. ABC narrative recording
   d. Conditional probability recording

   Hint: (See "Descriptive Functional Behavior Assessment")

10. Mr. Peterson has been struggling with Arnold, a boy with severe disabilities who hums and rocks back and forth intermittently throughout the day. Mr. Peterson wants to see if these behaviors are associated with any specific activities during the school day so that he can then more closely analyze what occurs during those time periods. Which descriptive assessment method would be the best choice for what he wants to accomplish?
    a. Scatterplot recording
    b. ABC continuous recording
    c. ABC narrative recording
    d. Functional analysis

    Hint: (See "Descriptive Functional Behavior Assessment")

11. Ms. Carmichael is conducting an FBA for Jamal, a boy in her class who runs away from activities and teachers. Ms. Carmichael began by interviewing her teaching assistants and by participating in the conversation with them to define the target behavior and to determine what antecedents and consequences she would watch for during her ABC assessment. Ms. Carmichael completes her ABC assessment and finds that running away from tasks and materials occurs most frequently when difficult tasks are presented and that the most common consequence for this behavior is a teacher chasing after him and returning him to the classroom. Sometimes he is required to return to the task and sometimes he is allowed to do a different activity when he comes back to the classroom. What can be concluded from this descriptive assessment?
    a. The problem behavior most likely occurs to get attention from teachers
    b. The problem behavior most likely occurs to escape from nonpreferred tasks.
    c. The problem behavior most likely occurs to get access to more preferred activities.
    d. The function of the problem behavior remains unclear.

    Hint: (See "Conducting a Functional Behavior Assessment")

12. After you complete a descriptive assessment, you should write hypothesis statements that reflect your interpretation of the data. Which of the following is a hypothesis statement that contains all of the important elements?
   a. Gain peer attention: When Valerie is playing alone on the playground, she bear hugs her peers.
   b. Gain peer attention: Valerie bear hugs her peers, which is followed by attention in the form of teasing from her peers.
   c. When Valerie is playing alone on the playground, she bear hugs her peers, which is followed by attention in the form of teasing from her peers.
   d. Gain peer attention: When Valerie is playing alone on the playground, she bear hugs her peers, which is followed by attention in the form of teasing from her peers.

   Hint: (See "Conducting a Functional Behavior Assessment")

13. The primary reason for conducting a functional analysis is:
   a. To test hypotheses generated via indirect and descriptive assessments.
   b. To generate hypotheses that can be further evaluated via indirect and descriptive assessments.
   c. To observe problem behavior within the naturally occurring routine.
   d. To identify any temporal patterns in problem behavior.

   Hint: (See "Overview of FBA Methods")

14. Characteristics of functional analysis include:
   a. They are conducted within naturally occurring routines.
   b. They are conducted within analog conditions that represent naturally occurring routines.
   c. They utilize interviews and rating scales rather than direct observation of problem behavior.
   d. They are conducted in clinical settings only.

   Hint: (See "Functional (Experimental) Analysis")

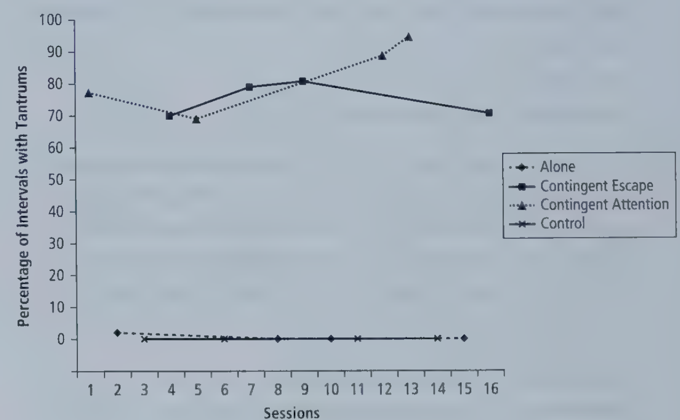
15. In addition to a control condition, a functional analysis typically consists of what test conditions?
   a. Contingent attention
   b. Contingent escape
   c. Alone
   d. All of these

   Hint: (See "Functional (Experimental) Analysis")

16. Mr. Moon is conducting a functional analysis with Ira, who kicks his teachers frequently. Mr. Moon is implementing a contingent escape condition. He prompts Ira to engage in a difficult task. Ira kicks Mr. Moon in the shin. How should Mr. Moon respond if he is to implement the contingent escape condition correctly?
   a. He should continue prompting Ira to do the task, but he should provide a mild reprimand, such as, "Ira, don't kick me. That hurts. Come on, it's time to work."
   b. He should remove the task materials and turn away from Ira for a short period of time.
   c. He should ignore Ira's kicking and continue with the task.
   d. He should provide Ira with a more preferred activity for a short period of time.

   Hint: (See "Table 27.1, page 682")

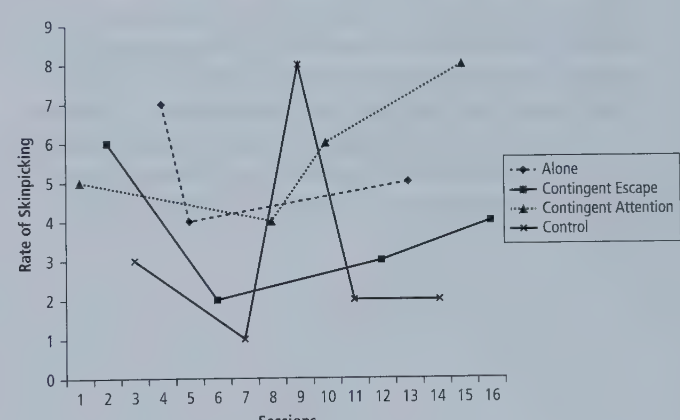
17. Look at the following graph from a functional analysis of Elsa's tantrums. What would you conclude is the function of Elsa's tantrums?



   a. Escape
   b. Attention
   c. Both escape and attention
   d. The pattern of behavior is undifferentiated; therefore the analysis is inconclusive.

   Hint: (See "Overview of FBA Methods")

18. Look at the following graph from a functional analysis of Walter's skin picking. What would you conclude is the function of Walter's skin picking?



   a. Escape
   b. Attention
   c. Automatic
   d. The pattern of behavior is undifferentiated; therefore the analysis is inconclusive.

   Hint: (See "Overview of FBA Methods")

ESSAY-TYPE QUESTIONS

1. Provide an example of a problem behavior that is maintained by social positive reinforcement, social negative reinforcement, and automatic reinforcement. In your answer, state what happens prior to the problem behavior, what the problem behavior is, and what follows the problem behavior.

   Hint: (See "Functions of Behavior")

2. Discuss three different ways in which the results of a functional behavior assessment can be used to formulate interventions that prevent future occurrences of problem behavior.

   Hint: (See "Role of Functional Behavior Assessment in Intervention and Prevention")

3. Explain the three types of functional behavior assessment. Discuss them in the order in which they might best be implemented and explain this sequence. Your explanation of the sequence should include the strengths and limitations of each type.

   Hint: (See "Overview of FBA Methods and Case Examples")

4. Compare and contrast descriptive and indirect assessments. What is the primary variable that sets these two forms of gathering descriptive information apart?

   Hint: (See "Overview of FBA Methods: Descriptive Functional Behavior Assessment and Indirect Functional Behavior Assessment")

5. Give an example of a descriptive assessment and state what specific information this assessment would yield.

   Hint: (See "Overview of FBA Methods: Descriptive Functional Behavior Assessment")

6. Assume you are a behavioral consultant for a local developmental disabilities agency. You have been asked to consult on young adult, Melissa, who is 22 years old and frequently disrobes while at work. You conduct indirect and descriptive assessments with her staff and learn the following: Melissa rarely disrobes when she is in the break room having a snack and interacting with her co-workers and staff. She disrobes most frequently when she is working on her assigned job task. It doesn't seem to matter what task she is working on. You conduct an ABC assessment and note that she works well when a staff member comes by to check her work, and this seems to continue for a few minutes after the staff member leaves. About 4 minutes after the staff member leaves, however, the clothes start coming off. Melissa is happy to put her clothes back on when directed to by a staff member. Because of these behaviors, Melissa's work station has been moved to behind a curtain in the work room. Much to the staff's chagrin, Melissa's disrobing has actually been worsening since they did this. What is your working hypothesis for Melissa's problem behavior? Write your hypothesis in the form recommended by the text, and provide a rationale for your hypothesis.

   Hint: (See "Conducting a Functional Behavior Assessment: Gathering Information and Interpreting Information and Forming Hypotheses")

7. Explain the difference between a functional analysis and other forms of functional behavior assessment (e.g., descriptive and indirect assessments). Specifically, why is it that a functional analysis allows you to test hypotheses, while other forms of functional behavior assessment are used for hypothesis formulation?

   Hint: (See "Overview of FBA Methods: Basic Procedure")

8. Describe the conditions that might be implemented in a functional analysis, as well as how and why each condition is implemented.

   Hint: (See "Overview of FBA Methods: Basic Procedure" and "Conducting a Functional Behavior Assessment: Testing Hypotheses")

9. Assume that in your role as a consultant, you decide to conduct a functional analysis of Melissa's disrobing (see Question 6). Assume that your functional analysis indicates her disrobing is maintained by both attention from others and escape from task demands. Assume that the functional analysis also confirmed your hypothesis that she rarely engages in problem behavior when she has attention and no task demands. Draw a graph that would illustrate this function appropriately.

   Hint: (See "Case Examples Illustrating the FBA Process")

PART 11

Special Applications

Parts 4 through 10 described basic principles of behavior and behavior change tactics derived from those principles. Part 11 describes several applications of behavior change technology. Each application represents a strategic approach to change behavior using multiple principles and procedures. Chapter 28 addresses the token economy, group contingencies, and contingency contracting. Chapter 29 details tactics individuals can implement to change their own behavior.

CHAPTER 28

Token Economy, Group Contingencies, and Contingency Contracting

LEARNING OBJECTIVES

- Define and identify the components of a contingency contract.
- Define and identify the components of a token economy.
- Explain the considerations that must be made when implementing a token economy.
- Define and identify the different types of group contingencies.
- Explain the considerations that must be made when implementing a group contingency.

This chapter addresses the token economy, group contingencies, and contingency contracting as behavior change strategies. Each strategy will be defined and its relationship to behavioral principles explained; requisite components of each strategy will be addressed and guidelines for designing, implementing, and evaluating each will be presented. These three strategies are grouped because they have several features in common. A well-established research base supports their effectiveness. Further, they can be combined with other approaches in package programs to provide an additive effect. Also, each strategy can be used with individual and multiple group arrangements. Finally, the flexibility of these three special applications makes them attractive options for practitioners in various settings.

TOKEN ECONOMY

The token economy is a highly developed and broadly researched behavior change system. It has been applied successfully in virtually every educational, residential, employment, and therapeutic setting possible (Bullock & Hackenberg, 2006; Doll, McLaughlin, & Barretto, 2013; Glynn, 1990; Hackenberg, 2009; Musser, Bray, Kehle, & Jenson, 2001; Pietras & Hackenberg, 2005; Stilitz, 2009; Wolfe, 1936). The utility of a token economy to change behaviors that previously have been resistant to modification is well established in the experimental and applied literatures. Further, token economies have proved successful whether administered by practitioners (Jowett Hirst, Dozier, & Payne, 2016) or even peer groups of children in elementary schools (Alstot, 2015). This section defines and describes the token economy and outlines effective procedures for using it in applied settings.

Token Economy Defined

A **token economy** (or *token reinforcement* system) consists of three major components: (a) a list of target behaviors; (b) tokens, points, or marks participants earn for emitting the target behaviors; and (c) a menu of **backup reinforcers**—preferred items, activities, or privileges—that participants obtain by exchanging earned tokens. Hackenberg (2009) described a token economy as "an interconnected set of contingencies that specifies the relations between token production, accumulation, and exchange" (p. 259). Operationally, these interconnections work in three stages. First, the behavior or behaviors to be reinforced are identified and defined. Second, a medium of exchange—the **token**—is selected. Third, participants exchange earned tokens for backup reinforcers (i.e., objects, privileges, or activities known or presumed to function as reinforcers for participants).

> In addition to having three stages, A token economy derives its complexity from the three interconnected schedules of reinforcement that govern when tokens will be delivered (token production), when tokens will be exchanged (exchange production), and the cost in tokens of the various goods or services available for exchange (token exchange; Hackenberg, 2009). These schedules of reinforcement represent internal mechanisms of a token economy that can be arranged using a number of basic reinforcement schedules (e.g., fixed ratio and variable interval) or complex schedules (e.g., differential reinforcement of other behavior). (Ivy, Meindl, Overley, & Robson, 2017, p. 3)

The Functions of Tokens

Store-based or manufacturer coupons function as tokens. When a customer purchases an item from a participating store, the cashier provides the purchaser with a paper coupon directly or electronically through a smartphone or tablet device. That coupon is traded later for another item at a reduced price, or it is redeemed immediately for the backup conditioned reinforcer.

Money, in the form of currency, is an example of a token that can be exchanged at a later time for backup objects and

activities (e.g., food, clothing, transportation, entertainment). As Bell and McDevitt (2014) stated:

> To see the enormous power of conditioned reinforcement, one must only consider the potent effects of money on human behavior. A "token" is an object that can be exchanged for other reinforcers. It is also called a "generalized reinforcer," as it does not need to rely on a single reinforcer in order to be effective. Money is one such token and maintains much of our behavior, but other kinds of tokens can be used to establish or maintain specific behaviors. (p. 242)

Hackenberg (2009) noted that, "A token, as the name implies, is an object of no intrinsic value; whatever function(s) a token has is established through relations to other reinforcers, both unconditioned (e.g., food or water) or conditioned (e.g., money or credit). As such, a token may serve multiple functions—reinforcing, punishing, discriminative, eliciting—depending on its relationship to these other events" (p. 259).

The multiple functions of tokens figure importantly in a token economy, which is often effective in applied settings, for three reasons. First, tokens bridge the time gap between the occurrence of a behavior and delivery of a backup reinforcer. For example, a token may be earned during the afternoon, but the backup reinforcer is not awarded until the next morning. Second, tokens bridge the setting gap between the behavior and the delivery of the backup reinforcer. For instance, tokens earned at school could be exchanged for reinforcers at home, or tokens earned in a general education classroom in the morning could be exchanged in a special education classroom in the afternoon. Finally, as generalized conditioned reinforcers, tokens make the management of motivation less critical for the behavior analyst.

Tokens are generalized conditioned reinforcers, and as such are independent of specific states of motivation because they are associated with a wide array of backup reinforcers. However, a generalized conditioned reinforcer is a relative concept: Effectiveness depends to a large extent on the extensiveness of the backup reinforcers and their ability to affect the future occurrence of behaviors (Moher, Gould, Hegg, & Mahoney, 2008). Tokens exchangeable for a wide variety of backup reinforcers have considerable utility in schools, clinics, centers, and hospitals where it is difficult to manage the delivery of powerful reinforcers, or deprivation states, which may change with time or location.

Higgins, Williams, and McLaughlin (2001), for example, used a token economy to decrease the disruptive behaviors of an elementary-age student with learning disabilities. The student exhibited high levels of out-of-seat, talking out, and poor sitting behaviors. After collecting baseline data on these behaviors, a token economy was implemented. If an alternative behavior was being emitted instead of the behaviors targeted for reduction, the student earned a checkmark exchangeable for free time. Maintenance checks were taken on two subsequent occasions to determine duration effects.

Figure 28.1 displays the results of the study across the three dependent variables. A functional relation was established between the onset of the token economy and the reduction of the behaviors. Further, maintenance checks indicated that the behavior remained at low levels after the token economy phase ended.

Level Systems

A **level system** is a type of token economy in which participants move up (and sometimes down) a hierarchy of tiers contingent on meeting specific performance criteria with respect to the target behaviors. As participants move "up" from one level to the next, they have access to more privileges and are expected to demonstrate more independence (see example in Table 28.1). The schedule of token reinforcement is gradually thinned so that participants at the highest levels are functioning on schedules of reinforcement similar to those in natural settings.

According to Smith and Farrell (1993), level systems are an outgrowth of two major educational advances that began in the late 1960s and 1970s: (a) Hewett's *Engineered Classroom* (1968) and (b) Phillips, Phillips, Fixen, and Wolf's *Achievement Place* (1971). In both cases, systematic academic and social programming combined token reinforcement, tutoring systems, student self-regulation, and managerial arrangements. Smith and Farrell stated that level systems are designed to

> foster a student's improvement through self-management, to develop personal responsibility for social, emotional, and academic performance . . . and to provide student transition to a less restrictive mainstream setting. . . . Students advance through the various levels as they show evidence of their achievement. (p. 252)

In a level system, students must acquire and achieve increasingly more refined repertoires while tokens, social praise, or other reinforcers are simultaneously decreased. Level systems have built-in mechanisms for participants to advance through a series of privileges, and they are based on at least three assumptions: (a) Combined techniques—so-called "package programs"—are more effective than individual contingencies introduced alone, (b) student behaviors and expectations must be stated explicitly, and (c) differential reinforcement is necessary to reinforce closer and closer approximations to the next level (Smith & Farrell, 1993).

Lyon and Lagarde (1997) proposed a three-level group of reinforcers that placed useful, but less desirable, reinforcers at Level 1. At this level, students had to earn 148 points, or 80%, of the 185 maximum points that could be earned during the week to purchase certain items. At Level 3, the powerful, and highly desirable, reinforcers could be purchased only if the students had accumulated at least 167 points, or 90%, of the total points possible. As the levels progress, the expectations for performance increase.

Cavalier, Ferretti, and Hodges (1997) incorporated a self-management approach with an existing level system to improve the academic and social behavior of two adolescent students with learning disabilities for whom increased participation in a general education classroom was an individualized education program (IEP) goal. Basically, other students in the classroom were making adequate progress through the six-level point system that the teacher had devised, but the inappropriate

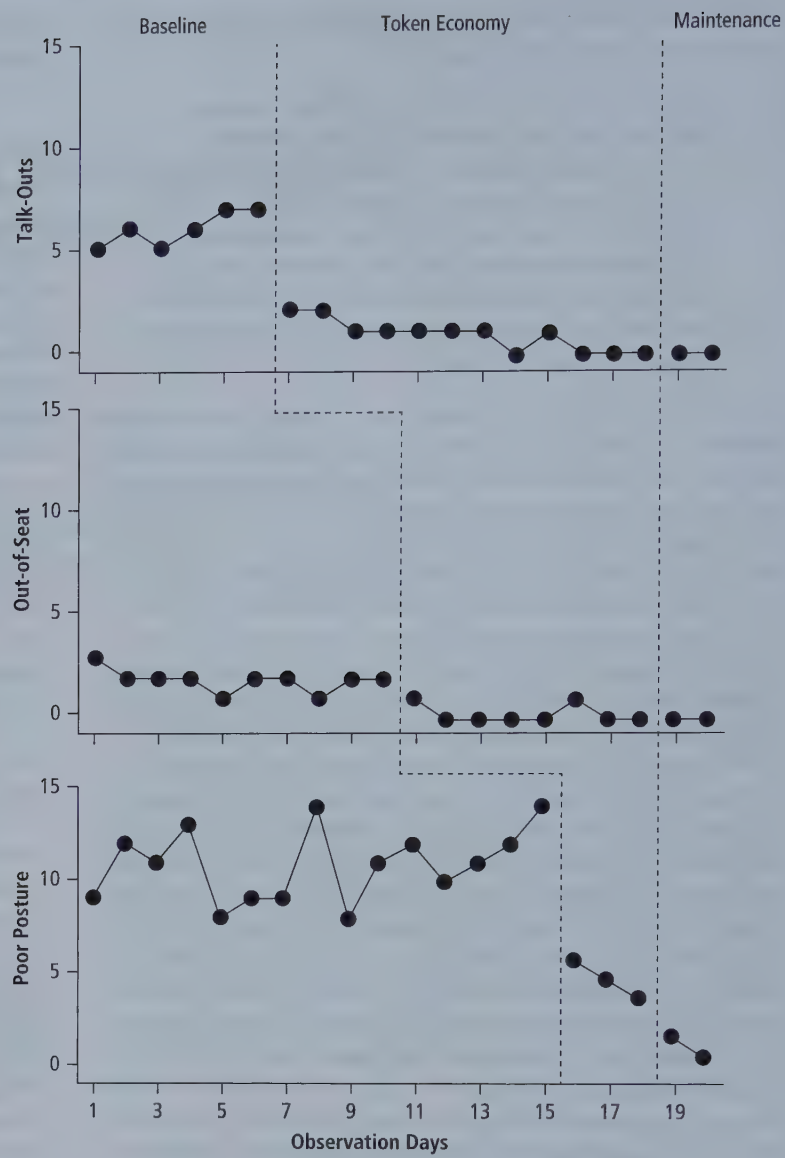


Figure 28.1 Number of talk-outs, out-of-seat behavior, and poor posture during baseline, token economy, and maintenance conditions.

From "The Effects of a Token Economy Employing Instructional Consequences for a Third-Grade Student with Learning Disabilities: A Data-Based Case Study," by J. W. Higgins, R. L. Williams, and T. F. McLaughlin, 2001, *Education and Treatment of Children*, *24*(1), p. 103. Copyright 2001 by The H. W. Wilson Company. Reprinted by permission.

verbalizations of the two target students during class stalled their progress at Level 1. After collecting baseline data on the number of inappropriate verbalizations, the researchers trained the students to self-record occurrences of these behaviors during two 50-min intervals in the day. Inappropriate verbalizations were defined explicitly, and during mock trials the students practiced self-recording and received feedback based on teacher observations during the same interval. An emphasis was placed on the accuracy of the students' recording, and desirable reinforcers were awarded for accurate recording. During the intervention (level system plus self-recording), students were told that they would be observed during the two 50-min intervals for accuracy. If they met the criterion for each level (e.g., a decrease of five inappropriate verbalizations from the previous session), a reinforcer was delivered. As the students progressed through the levels, the reinforcer changed to more highly desirable items.

Results showed a high number of verbalizations during baseline; however, when the packaged intervention was initiated for Student 1, inappropriate verbalizations decreased. This outcome was replicated for Student 2, confirming a functional relation between the intervention and the decrease in inappropriate verbalizations.

Designing a Token Economy

The basic steps in designing and preparing to implement a token economy are as follows:

1. Identify target behaviors and rules.
2. Select tokens.
3. Develop a menu of backup reinforcers.
4. Establish a token exchange ratio.
5. Determine when and how tokens will be dispensed and exchanged and what will happen if the requirements to earn a token are not met.
6. Field-test the system before full-scale implementation.

TABLE 28.1 Students' Increased Access to and Independent Use of Internet and Internet-enabled Devices as Rewards for Achieving Higher Levels of Performance

	Orientation phase	Growth phase		Independence phase	
Device	**Level 1**	**Level 1+**	**Level 2**	**Level 2+**	**Level 3**
Class computer	Internet access with staff supervision	Internet access with staff supervision	Internet access with staff supervision	Internet access with staff supervision	Internet access with staff supervision
Own games console	No	No	No	No	Internet access
Own handheld games console	No	No	Internet access with staff supervision	Internet access with staff supervision	Internet access
Houseroom computer	No	Internet access with staff supervision	Internet access with staff supervision	Internet access	Internet access
Community (café, library, etc.)	No	No	Internet access	Internet access	Internet access
Own laptop	No	No	Internet access with staff supervision	Internet access	Internet access
Own tablet computer	No	No	Internet access	Internet access	Internet access
Own basic mobile (calls and texts only)	No	No	No	Yes	Yes
Own smartphone	No	No	No	No	Internet access

From "The *ACHIEVE!* Program: A Point and Level System for Reducing Severe Problem Behavior," by D. Pritchard, H. Penney, and R. C. Mace, © 2018. Reproduce with permission of John Wiley & Sons Inc.

Identify Target Behaviors and Rules

Chapter 3 addressed the selection and definition of behavior change targets. The criteria presented in that chapter also apply to the selection and definition of rules and target behaviors for a token economy. Generally, the guidelines for selecting behaviors for a token economy include: (a) selecting only measurable and observable behaviors; (b) specifying criteria for successful task completion; (c) starting with a small number of behaviors, including some that are easy for the individual to accomplish; and (d) being sure the participants have the prerequisite skills for any targeted behaviors (Myles, Moran, Ormsbee, & Downing, 1992).

After rules and behaviors that apply to everyone are defined, then criteria and behaviors specific to individual learners should be established. Many token economy failures can be traced to requiring the same behaviors and setting the same criteria for all learners. Token economies usually need to be individualized. For example, in a classroom setting, the teacher may want to select different behaviors for each student, or perhaps the token economy should not be applied to all students in the classroom. Instead, only the students with the greatest challenges should be included. However, students who do not need a token system should still continue to receive other forms of reinforcement.

Select Tokens

Frequently used tokens include points, tally marks, plastic strips, checkers, coupons, poker chips, holes punched in a card, or teacher initials. Criteria to consider in selecting the token itself are important. First, tokens should be safe. If a young child or participant with severe learning or behavioral problems is to receive the token, it should not be an item that can be swallowed or otherwise used to cause injury. Also, tokens should not be large, cumbersome, or unwieldy. Second, the analyst should control token physical configuration; learners should not be able to bootleg the acquisition of tokens by surreptitiously manufacturing their own. If tally marks are used, they should be on a special card, or made with a special marking pen that is available only to the analyst. Likewise, if holes are punched in a card, the paper punch should be available only to the analyst to avoid counterfeiting.

Tokens should be durable because they may have to be used for an extended period, and they should be easy to carry, handle, bank, store, or accumulate. Tokens should also be readily accessible to the practitioner at the moment they are to be dispensed. It is important that they be provided immediately after the target behavior. Tokens should be inexpensive; there is no need to spend a large sum of money to purchase tokens. Injectable stamps, stars, checkmarks, and buttons are all inexpensive items that can be used repeatedly as tokens.

Although using desirable objects as tokens can backfire—a teacher who used baseball cards as tokens found students spent so much time interacting with the tokens (e.g., reading about players) that the tokens distracted students from the purpose of the system—for some students, the "objects of obsession" can serve as effective tokens. In a study by Charlop-Christy and Haymes (1998), three children with autism who attended an after-school program served as participants. All three children were consistently off task during activities, preoccupied with some objects, and engaged in self-stimulatory behaviors. During baseline, the students earned stars for appropriate behavior.

Inappropriate behaviors or incorrect responses were addressed by saying, "try again" or "no." When the students earned five stars, the tokens could be exchanged for backup reinforcers (e.g., food, pencils, tablets). During the token condition, an "obsession" object (one the children had been previously preoccupied with) was used as the token. Once they had earned five obsession objects, children could exchange them for food items or other known reinforcers. Charlop-Christy and Haymes reported that the overall pattern of responding demonstrated that when the token was an obsession object, student performance improved.

Develop a Menu of Backup Reinforcers

Most token economies can use naturally occurring activities and events as backup reinforcers. For example, in a classroom, center, or school setting, tokens can be exchanged for popular games or materials, or they can be traded for favorite classroom jobs, such as office messenger, materials distributor, teacher assistant, or media operator. Tokens can also be used for schoolwide privileges, such as a library or study hall pass; a special period (e.g., music, physical education) with another class; or special responsibilities, such as school patrol, cafeteria monitor, or tutor. Higgins and colleagues (2001) used naturally occurring activities and events as backup reinforcers for their token economy (i.e., the students had access to computer games and leisure books). However, play materials, hobby-type games, snacks, television time, allowance, permission to go home or downtown, sports events, and coupons for gifts or special clothing could also be used as backup reinforcers because these objects or items tend to occur in many settings.

If naturally occurring routine activities and events have not served as reinforcers in the past, then conducting a reinforcer assessment would be advisable (Sran & Borrero, 2010). Identifying backup reinforcers using a formal assessment procedure increases the chances that those stimulus items will serve as reinforcers. Further, an assessment would identify known reinforcers specific to the individual, and not suspected generic reinforcers for the group. Depending on the population of students to be served, analysts should be prepared to conduct multiple assessments over time as reinforcer preference can shift and change (DeLeon et al., 2001). All practitioners should use the most powerful, but least intrusive, and naturally occurring reinforcers possible.

Finally, selection of backup reinforcers should follow consideration of ethical and legal issues, as well as state and local education agency policies. Token reinforcement contingencies that would deny the learner basic needs (e.g., food) or access to personal or privileged information or events (e.g., access to mail, telephone privileges, attending religious services, medical care, scheduled recess) should not be used. Furthermore, general comforts that are associated with basic rights afforded to all citizens (e.g., clean clothing, adequate heating, ventilation, hot water) should not be used in a token program.

Establish a Token Exchange Ratio

Initially, the ratio between the number of tokens earned and the price of backup items should be small to provide immediate success for learners. Thereafter, the ratio of exchange should be adjusted to maintain the responsiveness of the participants. Following are general guidelines for establishing the ratio between earned tokens and the price of backup items:

- Keep initial ratios low.
- As token-earning behaviors and income increase, increase the cost of backup items, devalue tokens, and increase the number of backup items.
- With increased earnings, increase the number of luxury backup items.
- Increase the prices of necessary backup items more than those of luxury items.

Determine How the Tokens Will Be Dispensed

If objects such as tally marks or holes punched in a card are selected as the tokens, how the learner will receive them is obvious. If objects such as coupons or poker chips are used, some container should be made available for storing the accumulated tokens before they are exchanged for the backup items. Some practitioners have learners construct individual folders or containers for storing their tokens. Another suggestion is to deposit the tokens through cut slots in the plastic tops of coffee cans. With younger learners, tokens can be chained to form a necklace or bracelet.

Determine How the Tokens Will Be Exchanged

A menu of the backup items should be provided with a given cost for each item. Learners can then select from the menu. Some teachers have a table store with all the items displayed (e.g., games, balloons, toys, certificates for privileges). To avoid noise and confusion at shopping time, individual orders can be filled by writing in or checking off the items to be purchased. Those items are then placed in a bag with the order form stapled to the top and are returned to the purchaser. Initially, the store should be open frequently, perhaps twice per day. Lower functioning learners may need more frequent exchange periods. Later, exchange periods might be available only on Wednesdays and Fridays, or only on Fridays. As quickly as possible, token exchange should occur on an intermittent basis.

Specify What Happens if Token Requirements Are Not Met

Occasionally, token requirements will not be met for one reason or another. One approach is to nag the individual: "You didn't do your homework. You know your homework must be completed to earn tokens. Why didn't you do it?" A better approach is a matter-of-fact restatement of the contingency: "I'm sorry. You haven't enough tokens to exchange at this time. Try again." It is important to know whether the individual has the skills required to earn tokens. A learner should always be able to meet the response requirements.

What Should Be Done When a Learner Tests the System? How should a practitioner respond when a learner says she does not want any tokens or backup items? A bad approach would be to argue, debate, or cajole the learner. A better approach is to

say something neutral (e.g., "That is your decision.") and then walk away, precluding any argument or debate. In this way a confrontation is avoided, and the occasion remains set for token delivery for the learner. Most learners can and should have input in selecting the backup items, generating the rules for the economy, establishing the price for the backup items, and performing general duties in managing the system. A learner can be a salesperson for the store or a bookkeeper to record who has how many tokens and what items are purchased. When learners are involved and their responsibilities for the economy are emphasized, they are less likely to test the system.

Will the Token Economy Include a Response Cost Contingency? Procedures for including response cost—a negative punishment procedure—with a token economy were presented in Chapter 15. Token economies can include a token loss contingency for inappropriate behaviors and rule infractions (Musser et al., 2001). Any behaviors subject to response cost should be defined and stated clearly in the rules. Learners need to be aware of what actions will result in token loss and how much the behavior will cost. The more serious the inappropriate behavior is, the greater the token loss should be. Clearly, fighting, acting out, or cheating should result in greater token loss than minor infractions (e.g., out-of-seat behavior or talk-outs). Token loss should never be applied to a behavior if the learner does not have tokens. Students should not be allowed to go into debt, which would likely decrease the reinforcement value of the tokens. A learner should always earn more tokens than she loses.

Several studies have compared the effectiveness of token-loss (response cost) or token-earn contingencies. In response cost procedures, each participant starts with a number of tokens and loses a token on each instance of the problem behavior (e.g., Conyers et al., 2004; Donaldson, DeLeon, Fisher, & Kahng, 2014; McGoey & DuPaul, 2000). With a token-earn procedure, children earn tokens contingent on emitting specified desired behaviors or for the absence of problem behavior during specified intervals (DRO). Jowett Hirst et al. (2016) compared the effects of token-earn and token-loss contingencies on the on-task performance of students working in individual and group contexts. Results showed that both contingencies were effective in increasing on-task behavior (see Figure 8.17). When given a choice of which contingency would operate during each session, five of six children chose the response cost over the token earn.

Because the token-loss contingency required less time to implement, teachers who use a token system as a behavior management strategy might find arranging loss contingencies to be a better option than arranging earn contingencies. Box 28.1: "Coupons for Cursing," describes how a teacher incorporated response cost procedure into a classwide token economy.

Field-Test the System

The final step before actually implementing a token system is to field-test it. For 3 to 5 days, token delivery is tallied exactly as if tokens were being earned, but no tokens are awarded during the field test. Data from the field test are used for assessment. Are learners actually deficient in the targeted skills? Are some learners demonstrating mastery of behaviors targeted for intervention? Are some learners not receiving tokens? Based on answers to questions such as these, final adjustments in the system can be made. For some learners, more difficult behaviors may need to be defined; others may need less demanding target behaviors. Perhaps more or fewer tokens need to be delivered relative to the price of the backup reinforcers.

BOX 28.1

Coupons for Cursing

During my second year as a middle school teacher for students with learning disabilities and behavior disorders, I was plagued with a number of students for whom refraining from blurting out curse words and expletives was nearly impossible. At the first sign of frustration with a seatwork assignment or a negative encounter with a classmate, the expletives would roll out of their mouths like an "Old Faithful" gusher. Such outbursts were contagious and extremely problematic.

To address this issue, after collecting baseline data on the number of curses per morning, I devised a program that provided each target student with a stapled pack of 10 index cards on which I had written: **Coupon: Good for One Free Curse. Turn in After Cursing**.

I told the students in individual meetings with each of them that since our morning workload kept me busy, I would not have time to personally redirect them when they cursed. Instead, I told them that when I heard them curse, I would look at them and give them a personal signal (i.e., a tug on my ear, a tap to my nose, a touch to my shoulder, etc.). Tugs, taps, and touches were specific to each student. When the student saw that signal, he was to tear off the top coupon from the pack of 10 and put it on my desk. Also, he was told that all remaining coupons could be added to the pool of points that he otherwise would earn for completing academic work. In effect, the remaining coupons became "bonus tokens" exchangeable for inexpensive, but desirable, items that were available on Friday at the "Classroom Store." The program quickly reduced the number of expletives emitted by the target students. In effect, they learned that the extra coupons provided the opportunity to buy more highly desirable items that they otherwise would not have been able to afford.
—Timothy E. Heron

Implementing a Token Economy

Training Staff

To launch a token economy successfully, practitioners must learn all implementation elements. The behavior analyst can assist in this process by developing checklists of key operational components and teaching these components systematically using a direct instruction "model-lead-test" (MLT) procedure. The MLT procedure emphasizes the behavior analyst's role in demonstrating, role playing with trainees, conducting supervised and independent practice, and providing feedback and reinforcement. Further, and more importantly, the behavior analyst must recognize that during initial training, the trainees are on a learning curve and will make mistakes. The behavior analyst should be prepared to conduct booster training, review sessions, or monitoring protocols to ensure that the token system is being implemented as designed when the trainee is on her own.

Plavnick, Ferreri, and Maupin (2010) found that two 1-hr training sessions on how to implement the steps of a token economy did not produce high levels of implementation by three teachers in an early childhood special education class. Prior to training sessions, teachers implemented virtually none of the key elements of the token program. During training, the number of token steps executed correctly increased substantially for all three teachers. However, during the implementation phase, when teachers were conducting the program alone, the number of steps decreased markedly. The data show that the teachers failed to execute the checklist items they "learned" during training. When self-monitoring was introduced, the number of checklist items completed accurately increased and remained high. Plavnick et al. remind practitioners that additional support and monitoring may be needed after initial training to maintain high levels of performance.

Initial Token Training

The manner in which initial training is conducted to launch a token economy depends on the functioning level of the learners. For high-functioning learners, or those children with mild disabilities, initial training might require minimal time and effort and consist primarily of verbal instructions or modeling. Usually the initial token training for these individuals can be accomplished in one 30- to 60-min session. Three steps are normally sufficient. First, an example of the system should be given. The practitioner might describe the system as follows:

> This is a token and you can earn it by [specify behavior]. I will watch your behavior; and when you [specify behavior], you will earn a token. Also, as you continue [specify behavior], you will earn more tokens. At [specify time period] you will be able to exchange the tokens you have earned for whatever you want and can afford on this table. Each item is marked with the number of tokens needed for purchase. You can spend only the tokens you have earned. If you want an item that requires more tokens than you have earned, you will have to save your tokens over several [specify time period].

The second step is to model the procedure for token delivery. For instance, each learner, or all group members, might be directed to emit the specified behavior. Immediately following the occurrence of the behavior, the learner(s) should be praised (e.g., "Enrique, I'm pleased to see how well you are working by yourself!") and the token delivered.

The third step is to model the procedure for token exchange. Learners should be taken to the store and shown the items for purchase. All learners should already have one token, which was acquired during the modeling of token delivery. At this time, several items should be able to be purchased for one token (the cost may go up later)—a game, 5 minutes of free time, a sticker or teacher helper privilege. Students should actually use their tokens in this exchange. Lower functioning learners may require several sessions of initial token training before the system is functional for them. Further response prompts may be needed.

Ongoing Token Training

During token reinforcement training, the practitioner and the students should follow the guidelines for effective use of reinforcement (see Chapter 11). For example, tokens should be dispensed contingently and immediately after the occurrence of the desired behavior. Procedures for delivery and exchange should be clear and should be followed consistently. If a booster session is needed to improve teacher, parent, or student performance of how tokens are dispensed, earned, or exchanged, practitioners should deliver it. Finally, the focus should be on building and increasing desirable behaviors through token delivery rather than decreasing undesirable behaviors through response cost.

Management Issues During Implementation

Students must be taught how to manage the tokens they earn. For instance, once received, tokens should be placed in a safe, but accessible, container so that they are out of the way, but readily available when needed. If the tokens are in clear view and easily accessible, some students might play with them at the expense of performing academic tasks assigned by the teacher. Also, placing the tokens in a secure location reduces the risk of other students counterfeiting or confiscating them. Preemptive measures should be taken to ensure that tokens are not easily fabricated or reachable by anyone other than the recipient. If counterfeiting or stealing occurs, switching to different tokens will help reduce the likelihood of these tokens being exchanged under false pretenses.

Another management issue, however, relates to students' token inventories. Some students may hoard their tokens and not exchange them for backup reinforcers. Other students may try to exchange their tokens for a backup reinforcer, but they lack the requisite number of tokens to do so. Both extremes should be discouraged. That is, the program expectation should be that all students exchange at least some of their earned tokens periodically. Obviously, students without the requisite number of tokens should not be permitted to "borrow" against future projected earnings so as to participate in an exchange. Stated otherwise, they should not be permitted to buy backup reinforcers on credit.

A final management issue relates to chronic rule breakers, saboteurs, or students who test the system at every turn. Practitioners can minimize this situation by (a) ensuring that the tokens function as a generalized conditioned reinforcers, (b) conducting a reinforcer assessment to determine that the

backup items function as reinforcers and are preferred by each of the students, and (c) applying alternative procedures for chronic rule breakers. Specifically, if an individual student tries to sabotage the system, that student's program might be modified so that tokens are combined with a differential reinforcement of alternative behavior procedure, extinction, or brief time-out.

Withdrawing a Token Economy

Practitioners may be so encouraged by the effectiveness of the token economy they have worked so hard to design and implement that they do not want to remove the system. However, in most cases, practitioners and analysts should consider how they will gradually scale down, and ultimately remove, the program. One goal of the token program should be to have the descriptive verbal praise that is delivered simultaneously with the token acquire the reinforcing capability of the token. From the beginning, a systematic aim of the token economy should be to withdraw the program. Such an approach, aside from having functional utility for practitioners (i.e., they will not have to issue tokens forever), also has advantages for the learner. For example, if a special education teacher is using a token economy with a student scheduled for full-time placement in a regular fourth-grade classroom, the teacher wants to be certain that the student's responses can be maintained in the absence of the token economy.

Various methods have been used to withdraw token reinforcers gradually after criterion levels of behavior have been reached. The following six guidelines allow the practitioner to develop, and later withdraw, token reinforcers effectively.

First, the token presentation should always be paired with social approval and verbal praise. This should increase the reinforcing effect of the social approval and serve to maintain behaviors after token withdrawal.

Second, the number of responses required to earn a token should be gradually increased. For instance, if a student receives a token initially after reading only one page, he should be required to read more pages later for token delivery.

Third, the duration the token economy is in effect should be gradually decreased. For example, during September the system might be in effect all day; in October, the time might be 8:30 AM to 12:00 PM and 2:00 PM to 3:00 PM; and in November, the time could be reduced to 8:30 AM to 10:00 AM and 2:00 PM to 3:00 PM. In December, the times might be the same as those in November, but on only 4 days a week, and so on.

Fourth, the number of activities and privileges that serve as backup items and are likely to be found in the untrained setting should be increased gradually. For example, the analyst should start taking away tangible items in the store that might not be present in the regular classroom. Are edibles available at the store? They are usually not available for reinforcement in regular classrooms. Gradually, items should be introduced that would be common in a regular class (e.g., special award sheets, gold stars, positive notes sent home).

Fifth, the price of more desirable items should be increased systematically while keeping a very low price on less desirable items for exchange. For example, in a token system with adolescent girls who have moderate to severe intellectual disabilities, the prices of candy bars and trips to the canteen and grooming aids (e.g., comb, deodorant) were initially about the same. Slowly, the cost of items such as candy bars was increased to such a high level that the girls no longer saved tokens to purchase them. More girls used their tokens to purchase grooming aids, which cost substantially less than candy.

Sixth, the physical evidence of the token should be faded over time. The following sequence illustrates how the physical evidence of the token can be faded.

- Learners earn physical tokens such as poker chips or washers.
- The physical tokens are replaced with slips of paper.
- The slips of paper are replaced with tally marks on an index card that is kept by the learners.
- In a school setting, the index card can now be taped to the learner's desk.
- The index card is removed from the learner and is kept by the analyst, but participants can check their balance at any time.
- The analyst keeps tallies, but no checking is allowed during the day. Totals are announced at the end of the day, and then every other day.
- The token system is no longer operative. The behavior analyst does not announce point totals even though they are still kept.

Evaluating a Token Economy

Token economies can be evaluated using any number of reliable, valid, field-tested, best-practice designs (e.g., reversal, multielement). Given that most token economy programs are conducted with small groups, single-subject evaluation designs allow for participants to serve as their own control. Applied behavior analysts should always strive to evaluate effectiveness at the individual level because despite aggregate group data demonstrating effectiveness, individual data may not be so robust (Davis & Chittum, 1994).

The Jowett Hirst et al. (2016) study described earlier is informative here. When group data were aggregated, both procedures were shown to be equally effective. However, when individual data served as the basis of comparison, many students performed better under the response cost contingency (see Figure 8.17). Had an individual analysis not been conducted, these results would have been masked by the overall group data. The authors concluded

> . . . the procedures were differentially effective for some individuals in the group, which suggests that analyzing individual data is important because these differences may not have been observed if we reported only group averages. The importance of analyzing individual data is further supported by the results of Study 2, which showed differential effects for three participants . . . , whereas the overall results suggest that the two procedures are equally effective. (pp. 340–341)

Collecting social validity data before, during, and/or after token economy with the target participants and the significant others helps to answer the question "Did the result(s) of token implementation make a difference for the participant(s)?" Jowett Hirst et al.'s (2016) post-study preference assessment showed

that the majority of students preferred the response cost condition, an interesting finding given the negative punishment protocol associated with this procedure. Importantly, such a finding gives at least slight pause to the generally recommended guidance to use "positive" approaches initially to effect behavior change, when evidence to the contrary shows that for some students, a negative punishment procedure may be effective and preferred.

Further Considerations

Its effectiveness notwithstanding, the token economy does present some challenges to learners and practitioners.

Intrusive. Token systems can be intrusive. It takes time, energy, and resources to establish, implement, and evaluate token programs. Also, because most natural environments do not reinforce a person's behavior with tokens, careful thought must be given to how to thin the token schedule while simultaneously maintaining performance. In any case, token economy programs can have lots of "moving parts," and practitioners must be prepared to deal with them.

Self-perpetuating. A token economy can be an effective procedure for managing behavior, and analysts can be so encouraged by the results that they do not want to remove the system. Learners then continue working for backup reinforcers that may not normally be available in the natural environment.

Cumbersome. Token economies can be cumbersome to implement, especially if there are multiple participants with multiple schedules of reinforcement. The system may require additional time and effort from the learner and the behavior analyst.

However, even though the effectiveness of the token economy has been shown repeatedly, Hackenberg (2018) states there is much still to learn:

> Work in the area of token systems is best served by a bidirectional interplay between laboratory and applied research, where applied questions inspire research on basic mechanisms. When based on and contributing to an analysis, applied research on token economies can be on the leading edge of theoretical advances, helping set the scientific research agenda. (393)

GROUP CONTINGENCIES

Thus far in this text we have focused primarily on how contingencies of reinforcement can be applied to individuals. However, applied research has also demonstrated how contingencies can be applied to groups. Behavior analysts have increasingly turned their attention toward group contingencies as interventions in such areas as decreasing disruptive behavior in public school settings (Stage & Quiroz, 1997), increasing leisure activities of adults with traumatic brain injury (Davis & Chittum, 1994), promoting abstinence from cigarette smoking (Meredith, Grabinski, & Dallery, 2011), improving behavior and academic performance of children in schools and classrooms (Skinner, Skinner, Skinner, & Cashwell, 1999; Popkin & Skinner, 2003; Skinner, Cashwell, & Skinner, 2000; Skinner, Skinner & Burton, 2009), and reducing problem behaviors on playgrounds (Lewis, Powers, Kelk, & Newcomer, 2002), to name just a few. Each of these applications has shown that group contingencies, properly managed, can be an effective and practical approach to changing the behavior of many people simultaneously (Stage & Quiroz, 1997). Longitudinal data collected over a 30-year period support the proposition that group contingencies with general education and exceptional children populations, applied across a myriad of settings, have been effective (Little, Akin-Little, & O'Neill, 2015).

Group Contingency Defined

In a **group contingency** a common consequence (usually, but not necessarily, a reward intended to function as reinforcement) is contingent on the behavior of one member of the group, the behavior of part of the group, or the behavior of everyone in the group. Group contingencies can be classified as dependent, independent, or interdependent (Litow & Pumroy, 1975).

Rationale for and Advantages of Group Contingencies

There are a number of reasons for using a group contingency in applied settings. First, it can save time during administration. Instead of repeatedly administering a consequence to each member of a group, the practitioner can apply one consequence to all members of the group. From a logistical perspective, a practitioner's workload may be reduced. A group contingency can be effective and economical, requiring fewer practitioners or less time to implement.

Another advantage is that a practitioner can use a group contingency in a situation in which an individual contingency is impractical (Hanley & Tiger, 2011). For example, a teacher attempting to reduce disruptive behaviors of several students might have difficulty administering an individual program for each pupil in the classroom. A substitute teacher, in particular, might find the use of a group contingency a practical alternative because her knowledge of the students' previous histories of reinforcement would be limited and the group contingency could be applied across a variety of behaviors, settings, or students.

A group contingency can also be used in cases in which the practitioner must resolve a problem quickly, as when serious disruptive behavior occurs. The practitioner might be interested not only in decreasing the disruptive behavior rapidly but also in building improved levels of appropriate behavior.

Furthermore, a group contingency can capitalize on peer influence or peer monitoring because this type of contingency sets the occasion for peers to act as change agents (Gable, Arllen, & Hendrickson, 1994; Skinner et al., 1999). Admittedly, peer pressure can have a detrimental effect on some people; they may become scapegoats, and negative effects may surface (Romeo, 1998). However, potentially harmful or negative outcomes can be minimized by structuring the contingency elements randomly (Kelshaw-Levering, Sterling-Turner, Henry, & Skinner, 2000; Popkin & Skinner, 2003).

Finally, practitioners can establish a group contingency to facilitate positive social interactions and positive behavioral supports within the group (Kohler, Strain, Maretsky, & DeCesare, 1990). For example, a teacher might establish a group contingency

for a student or a group of students with disabilities. The students with disabilities might be integrated into the general education classroom, and a contingency could be arranged in such a way that the class would be awarded free time contingent on the performance of one or more of the students with disabilities.

Independent Group Contingencies

An **independent group contingency** is an arrangement in which a contingency is presented to all members of a group, but reinforcement is delivered only to those group members who meet the criterion outlined in the contingency (see Figure 28.2). Independent group contingencies are frequently combined with contingency contracting and token reinforcement programs because these programs usually establish reinforcement schedules independent of the performance of other members of the group.

Brantley and Webster (1993) used an independent group contingency in a general education classroom to decrease the disruptive behavior of 25 fourth-grade students. After collecting data on off-task behavior, call-outs, and out-of-seat behavior, the teachers posted rules related to paying attention, seeking the teachers' permission before talking, and remaining in their seats. An independent group contingency was established whereby each student could earn a checkmark next to his or her name on a list that was posted publicly in the room during any of the intervals that marked the observation periods during the day. When a student emitted an appropriate or prosocial behavior, a checkmark was registered. The criterion for earning a reward was increased from four to six checkmarks over 4 out of 5 days per week.

Results showed that after 8 weeks, the total number of combined disruptions (e.g., off-task behavior, call-outs, and out-of-seat behavior) decreased by over 70%, and some off-task behaviors (e.g., not keeping hands to self) were eliminated completely. The teacher's satisfaction with the approach was positive, and parents reported that they were able to understand the procedures that were in place for their children at school. Brantley and Webster (1993) concluded:

> The independent contingency added structure for students by using clear time intervals and clarified teacher expectations by limiting and operationally defining rules to be followed, monitoring behavior consistently, and setting attainable criteria for students. (p. 65)

A →	B →	C
Criterion stated for all group members*	Criterion met	S^{R+} for *each* member achieving criterion

*(E.g., "Each student who spells 9 out of 10 words correctly on the test Friday will earn 10 bonus points.")

Figure 28.2 An independent group contingency.

From Hirst, E. S. J., Dozier, C. L., Payne, S. W. (2016). Efficacy of and preference for reinforcement and response cost in token economies. *Journal of Applied Behavior Analysis, 49*, 329–345. Reproduced with permission of John Wiley & Sons,Inc.

A →	B →	C
Criterion stated for one individual or small group*	Criterion met	S^{R+} for whole group

*(E.g., "When all students at Table 2 finish their math assignments, the class will have 5 minutes of free time.")

Figure 28.3 A dependent group contingency.

From "The Effects of Self-Monitoring on the Procedural Integrity of a Behavioral Intervention for Young Children with Developmental Disabilities" by J. B., Plavnick, S. J. Ferreri, S. J., and A. N. Maupin, 2010, *Journal of Applied Behavior Analysis, 43*, 318. Reproduced with permission of John Wiley & Sons,Inc.

Dependent Group Contingencies

Under a **dependent group contingency**, the reward for the whole group is dependent on the performance of an individual student or small group. Figure 28.3 illustrates the dependent group contingency as a three-term contingency. The contingency operates like this: If an individual (or small group within the total group) performs a behavior to a specific criterion, the group shares the reinforcer. The group's access to the reward depends on the individual's (or small group's) performance. If the individual performs below the criterion, the reward is not delivered. When an individual, or small group, earns a reward for a class, the contingency is sometimes referred to as the **hero procedure**. According to Kerr and Nelson (2002), the hero procedure can facilitate positive interactions among students because the class as a whole benefits from the improved behavior of the student targeted for the group contingency.

Gresham (1983) used an interesting variation of the hero procedure within a dependent group contingency whereby the contingency was applied at home, but the reward was delivered at school. In this study, an 8-year-old boy who was highly destructive at home (e.g., set fires, destroyed furniture) earned good notes for nondestructive behavior at home. Billy received a good note—a daily report card—each day that no destructive acts took place. Each note was exchangeable for juice, recess, and five tokens at school the next day. After Billy received five good notes, the whole class received a party, and Billy served as the host. Gresham reported that the dependent group contingency reduced the number of destructive behaviors at home, and represented the first application of a dependent group contingency in a combined home–school setting.

Interdependent Group Contingencies

An **interdependent group contingency** is one in which all members of a group must meet the criterion of the contingency (individually *and* as a group) before any member earns the reward (Elliot, Busse, & Shapiro, 1999; Kelshaw-Levering et al., 2000; Lewis et al., 2002; Skinner et al., 1999; Skinner et al., 2000) (see Figure 28.4). Theoretically, interdependent group contingencies have a value-added advantage over dependent and independent group contingencies insofar as they yoke students to achieve a common goal, thereby capitalizing on peer pressure and group cohesiveness.

A ⟶ B ⟶ C

Criterion stated for all group members* | Criterion met | S^{R+} for whole group *if* all members of the group achieve criterion

*(E.g., "Each student must complete at least four science projects by the 6th week in the term in order for the class to go on the field trip.")

Figure 28.4 An interdependent group contingency.

The effectiveness of dependent and interdependent group contingencies may be enhanced by randomly arranging some or all components of the contingency (Popkin & Skinner, 2003). That is, randomly selected students, behaviors, or reinforcers are targeted for the contingency (Kelshaw-Levering et al., 2000; Skinner et al., 1999). Kelshaw-Levering and colleagues (2000) demonstrated that randomizing either the reward alone or multiple components of the contingency (e.g., students, behaviors, or reinforcers) was effective in reducing disruptive behavior.

Procedurally, an interdependent group contingency can be delivered (a) when the group as a whole meets the criterion, (b) when the group achieves a mean group score (Kuhl, Rudrud, Witts, & Schulze, 2015), or (c) based on the results of the Good Behavior Game or the Good Student Game. In any case, interdependent group contingencies represent an "all or none" arrangement. That is, all students earn the reward or none of them do (Popkin & Skinner, 2003).

Total Group Meets Criterion

Lewis and colleagues (2002) used the *total group meets criterion* variation to reduce the problem playground behaviors of students enrolled in a suburban elementary school. After a faculty team conducted an assessment of problematic playground behaviors, social skill instruction in the classroom and on the playground was coupled with a group contingency. During social skills instruction, students learned how to get along with friends, cooperate with each other, and be kind. During the group contingency, students earned elastic loops they could affix to their wrists. After recess, students placed the loops in a can on the teacher's desk. When the can was full, the group earned a reinforcer. Figure 28.5 shows the results of the social skills plus group contingency intervention across three recess periods during the day.

Group Averaging

Baer and Richards (1980) used a group averaging interdependent group contingency to improve the math and English performance of 5 elementary-aged students. In their study, all of the students in a class of 10, including the 5 target students, were told they would earn 1 extra minute of recess for each point of class improvement beyond the previous weekly average. Also, all students were given a contract stating this same contingency. The extra recess was awarded every day of the following week. For example, if the students' weekly averages exceeded their previous weekly averages by 3 points, they would receive 3 minutes of extra recess every day of the following week. The results of this 22-week study showed that all students improved when the group contingency was in effect.

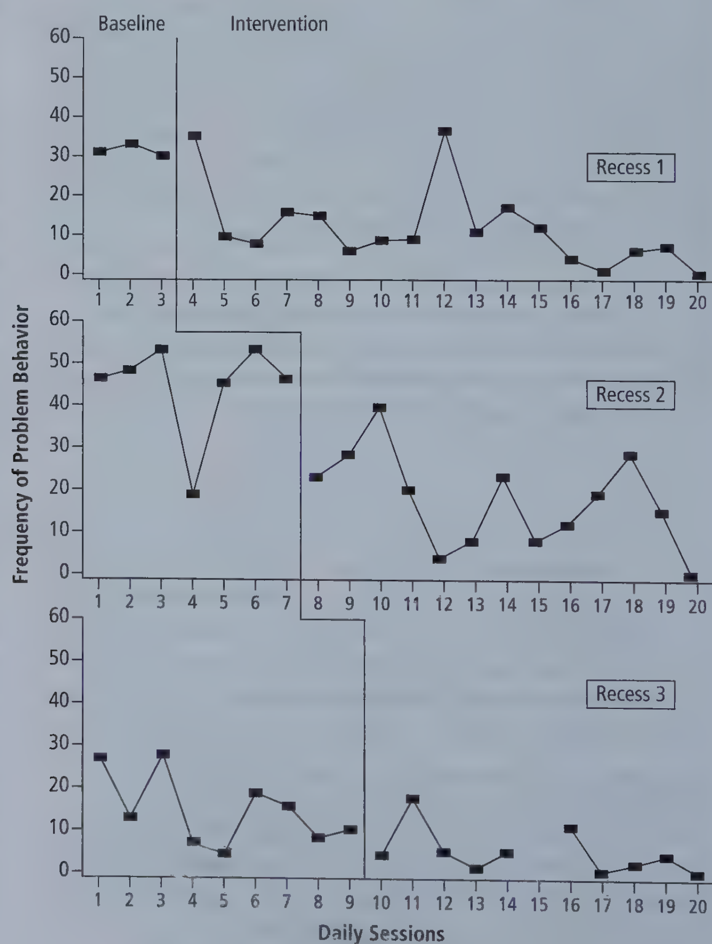


Figure 28.5 Frequency of problem behaviors across recess periods. Recess 1 was composed of second- and fourth-grade students, Recess 2 was composed of first- and third-grade students, and Recess 3 was composed of fifth- and sixth-grade students. Kindergarten students were on the playground across Recess 1 and 2.

From "Reducing Problem Behaviors on the Playground: An Investigation of the Application of School-Wide Positive Behavior Supports" by T. J. Lewis, L. J. Powers, M. J. Kelk, and L. L. Newcomer, © 2002. Reproduce with permission of John Wiley & Sons Inc.

Kuhl et al. (2015) used an interdependent group contingency as part of a comparison study on the effects of group versus individual goal setting on the increased level of physical activity of 4 selected students (2 males and 2 females) across two classrooms with total class populations of 25 to 30 students, respectively. In their study, students received a pedometer that registered their total number of steps during a 24-hour period, 4 days per week (Monday through Thursday). Prior to launching the study, teachers selected five activities from a list of 20 potential items thought to be reinforcing to the students. Ultimately, from that list of five activities, students voted for 15 minutes of extra free time as their reinforcer. Recess was awarded on Fridays.

A counterbalanced reversal design (ABACX) was used where X served as the best phase. During baseline, students received a calibrated pedometer, set to zero steps, which the students wore for a 24-hour period, excluding time spent bathing/swimming, or sleeping, over the course of the study. During baseline, students could not see the number of steps taken, as the pedometer was masked. Data were recorded daily, 4 days per week. During the Classroom Cumulative Total Condition, a cumulative total goal was determined by averaging each participant's daily step count, and then adding 1500 steps per student per day to that average.[1] Fifteen minutes of extra free time was awarded on Friday if the cumulative total of steps exceeded the goal. Pedometers were unmasked in this condition. Each day, the teachers publically posted cumulative daily step counts and provided verbal feedback to show progress toward their goal. When the Individual Goals Condition was introduced, the children were given their unmasked pedometers, but they met with the teacher, who informed them of their individual goal, which was determined by averaging the previous 4 days of data and adding 1500 steps to that figure. If a student met his goal, he was praised. If the goal was not met, the teacher encouraged the student to keep trying to meet the goal. Extra free time was administered on Friday if 80% of the students met their individual goal each day. Finally, a feedback schedule-thinning phase was introduced. During the 2-day condition, step count and recording occurred on only 2 days. During the 4-day condition, counts and recordings occurred on only 1 day per week. Reinforcement conditions remained the same.

The results of the study (see Figure 28.6) showed that step counts increased under both the cumulative and individual

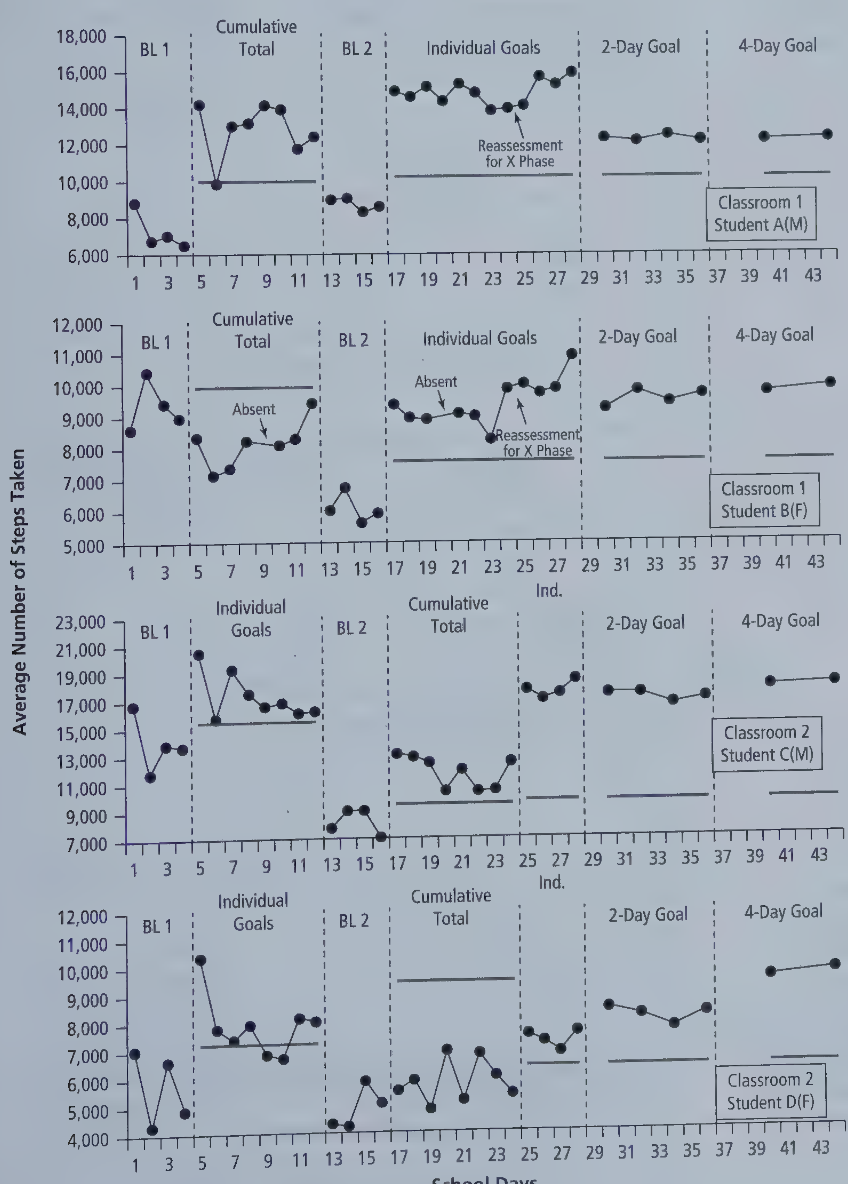


Figure 28.6 Total number of steps by four randomly selected students during cumulative and individual goal setting by four students.

From "Classroom-based Interdependent Group Contingencies Increase Children's Physical Activity," by S. Kuhl, E. H., Rudrud, B. N. Witts, and K. A. Schulze, 2015, *Journal of Applied Behavior Analysis, 48*, p. 609. Reproduced with permission of John Wiley & Sons,Inc.

contingencies; however, higher step counts were registered under the individual contingency regardless of its order during counterbalancing. Anecdotal evidence, collected after the study, suggested that under the individual contingency students engaged in extra walking practices that otherwise may not have occurred. That is, they walked in the gym before school, played action games with peers after school, and had parents drop them off farther from school each day to increase step counts. The fact that step counts during the thinning phase were still higher than baseline, but lower than during the individual phase, suggested the important role of daily feedback in establishing and maintaining increased step counts.

Good Behavior Game

Barrish, Saunders, and Wolf (1969) first investigated the **Good Behavior Game** (GBG) as an interdependent group contingency to address a variety of student classroom disruptive behaviors. Procedurally, to play the GBG a class (or group) is divided into two or more teams. Prior to the game being played, the teams are told that whichever team has the *fewest* marks against it at the end of the game will earn a reward. Each team is also told that it can win a reward if it has fewer than a specified number of marks (a differential reinforcement of low rates [DRL] schedule). Data reported by Barrish et al. show that this strategy can be an effective method of reducing disruptive behavior in the classroom. When game conditions were in effect during math or reading, talking-out and out-of-seat behaviors occurred at low levels. When game conditions were not in effect, disruptive behaviors occurred at much higher levels (see Figure 28.7).

In the Good Behavior Game, teacher attention is directed toward observing and recording occurrences of misbehavior, with the incentive that if one or more teams have fewer than the criterion number of infractions, a reinforcer is delivered. The advantage of the Good Behavior Game is that competition can occur within teams, across teams, or against a criterion.

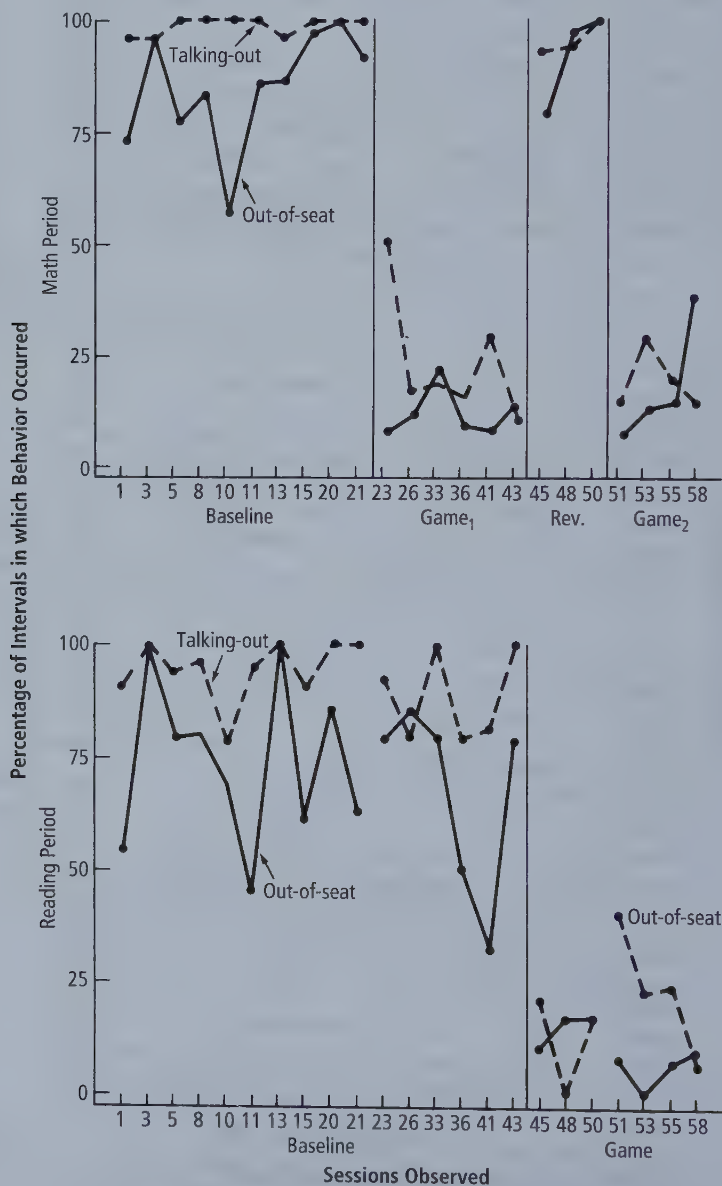


Figure 28.7 Percentage of 1-min intervals containing talking-out and out-of-seat behaviors in a classroom of 24 fourth-grade children during math and reading periods.

From "Good Behavior Game: Effects of Individual Contingencies for Group Consequences on Disruptive Behavior in a Classroom" by H. H. Barrish, M. Saunders, and M. M. Wolf, 1969, *Journal of Applied Behavior Analysis*, *2*, p. 122. Reproduced with permission of John Wiley & Sons,Inc.

Since Barrish et al.'s groundbreaking study, a number of researchers have investigated the GBG across a broad spectrum of settings and age groups, and in combination with other procedures (for reviews, see Little, Akin-Little, & O'Neill, 2015; Tingstrom, Sterling-Tuner, & Wilczynski, 2006). Further, several investigators have demonstrated its effectiveness as a preventative measure (Embry, 2002; Lannie & McCurdy, 2007; Van Lier, Van Der Sar, Muthen, & Crijen, 2004). By implementing a successful GBG program with children who are young—but disruptive and out-of-control—the goal is to "prevent" even worse outbursts in middle or high school by presumably socializing students at an early age to expectations, rules, and contingencies of reinforcement.

Donaldson, Vollmer, Krous, Downs, and Berard (2011), for example, conducted a replication of the Good Behavior Game with five teachers across their respective kindergarten classes, each experiencing disruptive students during morning group work that interfered with the goals of the lesson. After collecting baseline data on the total rate of disruptive behaviors, the class was divided into two teams, each containing, based on the teacher's judgment, a comparable number of disruptive students. During the GBG, teachers provided rules related to the children staying on their carpet patches, keeping their hands and feet to themselves, and speaking when called upon. If a disruptive behavior occurred, the experimenter—and later the teacher—placed a mark on the board and simultaneously announced aloud what rule was broken by a member of that team. One team won the game by having fewer points than the other team, or both teams won if they both met a criterion set by the teacher of an 80% or better reduction from baseline levels of misbehavior.

Figure 28.8 shows that under the GBG contingency, disruptive behavior decreased across all classrooms. Further, low levels of disruptive behavior continued after the experimenter shifted game implementation to the teacher, despite treatment integrity data indicating that teachers did not implement all procedures as trained. In short, even though implementation steps were likely compromised due to teachers simultaneously conducting an academic lesson, results were still robust.

Finally, teachers reported anecdotally that the children, when offered a choice, "voted" to continue playing the game for the rest of the year, and four teachers continued the game the following year. Donaldson et al. (2011) suggest that the (a) rewards that the students earned for winning may have served as social reinforcers, (b) hash marks on the board for disruptive behavior may have served as punishers, or (c) children's social interplay during the game (or later) may have affected the overall outcome. In any case, this study showed the powerful effects of the GBG with a very young population of students.

Kleinman and Saigh (2011) also conducted a successful replication study of the GBG with a class of extremely disruptive high school–level students attending a multiethnic public high school in a large urban area (Harlem, New York). Students were

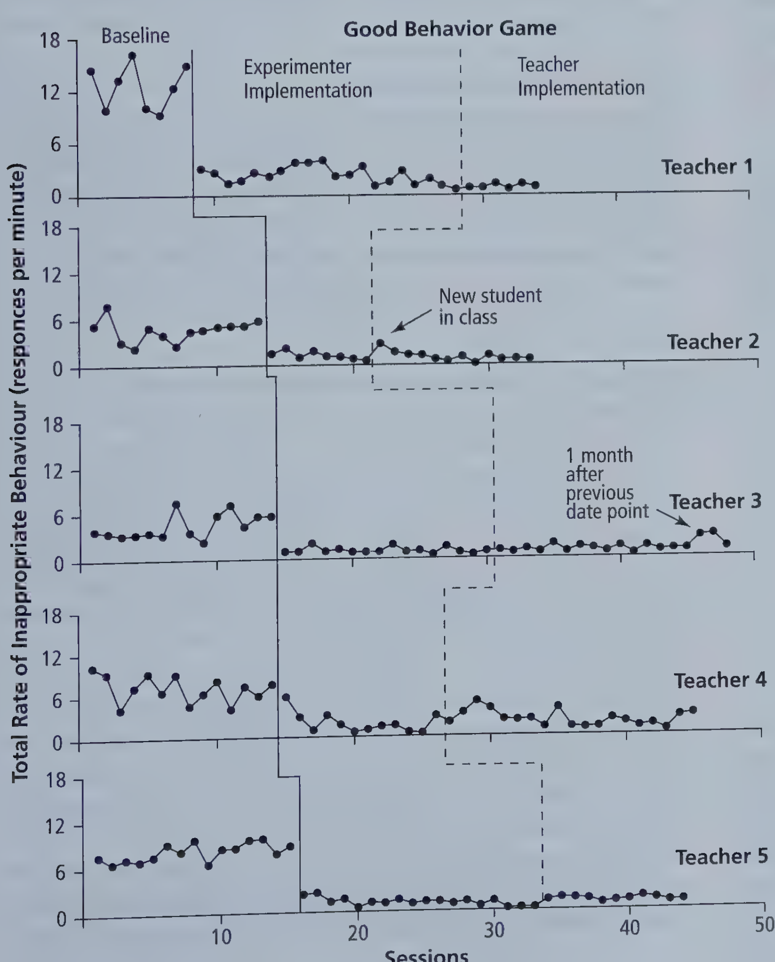


Figure 28.8 Rate of inappropriate behavior by kindergarten students as a function of the Good Behavior Game implemented by the experimenter and classroom teachers.

From "An Evaluation of the Good Behavior Game in Kindergarten Classrooms," by J. M. Donaldson, T. R. Vollmer, T. Krous, S. Downs, and K. P. Berard, 2011, *Journal of Applied Behavior Analysis*, *44*, pp. 605–609. Reproduced with permission of John Wiley & Sons,Inc.

observed to be so highly disorderly, with out-of-seat behavior, calling out, and physical altercations in class, that a previous teacher resigned her position. The new teacher's reception was met with cursing, screaming, and throwing objects. To address this situation, the new teacher conducted a reinforcer preference phase that was used to determine rewards that could be delivered given the ages of the students and budget constraints imposed by circumstances. A reversal design was employed that involved a short adaptation period to acclimate the students to observers in the class, followed by a Baseline I, GBG, Baseline II, and GBG 2.

During Baseline I, students were divided into two teams with approximately equal numbers of disruptors on each team. Further, the teacher posted a list of "expectations" for student behavior on the front board and read them before each session. Finally, all disruptions were handled in the past manner (e.g., teacher reprimands, removal from the classroom). During GBG 1, the teacher announced to the students that the teams would now compete with each other for daily and weekly prizes. They were told that any occurrence of defined target behaviors would result in the disruptor being named aloud and a mark placed against that disruptor's team. At the end of the period, the team with fewer marks would earn a prize (e.g., bite-sized candy). A weekly prize (i.e., a pizza party, cupcakes) would be awarded to the team with fewer marks for the week. Baseline II conditions replicated Baseline I. GBG 2 replicated GBG 1. A follow-up phase was conducted several weeks after the study's completion. Results indicated that when baseline conditions were in effect, extremely high levels of disruptive behavior occurred. When GBG conditions were in effect, all levels of disruptive behavior decreased markedly, and remained so during a successive follow-up phase (see Figure 28.9).

Kleinman and Saigh concluded that the GBG (a) reduced the disruptive target behaviors rapidly, (b) provided strong external validity of the efficacy of the GBG for older students, and (c) produced convincing social validity approval data from students in an inner-city, multiethnic setting. That is, the students themselves reported that they could observe the dramatic effects of improved behavior during GBG phases.

Good Student Game

The *Good Student Game* combines an interdependent group contingency (like the Good Behavior Game) with an added self-monitoring twist (Babyak, Luze, & Kamps, 2000). Basically, the Good Student Game is intended for implementation during independent seatwork periods when problematic behaviors surface. In the Good Student Game, the teacher (a) chooses target behaviors to modify, (b) determines goals and rewards, and (c) determines whether group or individual monitoring (or both) will occur.

Students are trained in the Good Student Game using a model-lead-test instructional sequence: Students are arranged in clusters of four or five, the target behaviors are defined, examples and nonexamples are provided, practice occurs under teacher supervision, and one or more students record their own or the group's performance. Table 28.2 shows how the Good Student Game compares to the Good Behavior Game. Note that two distinctions relate to target behaviors, the delivery of rewards, and feedback.

Guidelines for Implementing a Group Contingency

Implementing a group contingency requires as much preplanning as any other behavior change procedure. Presented here are six guidelines to follow before and during the application of a group contingency.

Choose an Effective Reward

One of the more important aspects of a group contingency is the strength of the consequence; it must be strong enough to serve as an effective reward. Practitioners are advised to use generalized conditioned reinforcers or reinforcer menus at every opportunity. Both of these strategies individualize the contingency, thereby increasing its power, flexibility, and applicability.

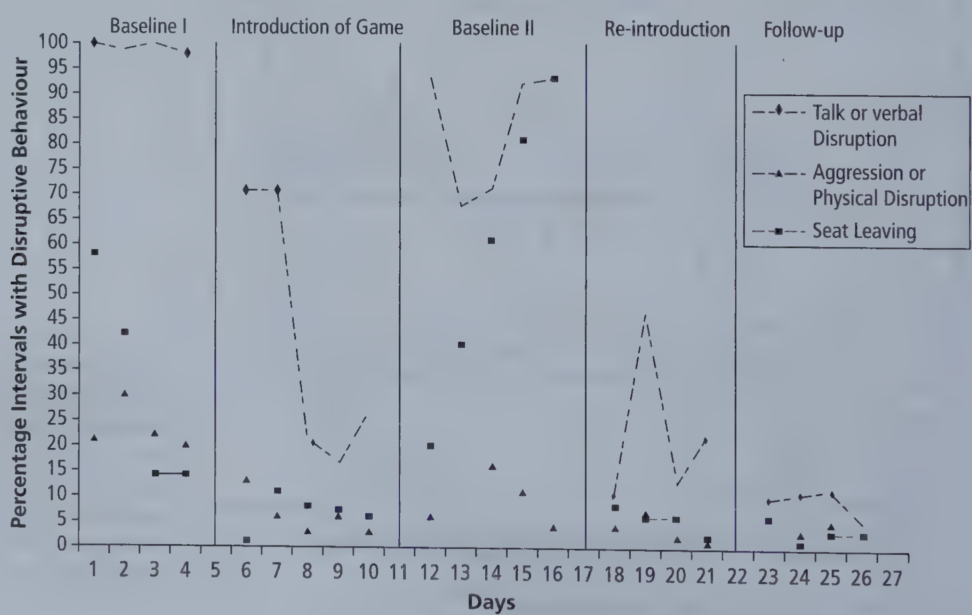


Figure 28.9 Disruptive behavior by general education high school students as a function of the Good Behavior Game.

From "The Effects of the Good Behavior Game on the Conduct of Regular Education New York City High School Students," by K. E. Kleinman and P. A. Saigh, 2011, *Behavior Modification, 35*(1), p. 101. Copyright 2011 by Sage Publications. Reprinted by permission.

TABLE 28.2 Components of the Good Behavior Game and the Good Student Game

Component	Good Behavior Game	Good Student Game
Organization	Students play in teams.	Students play in teams or as individuals.
Management	Teacher monitors and records behavior.	Students self-monitor and record their behavior.
Target behaviors	Behaviors are stated as rule breaking or rule following.	Behaviors are stated as rule following.
Recording	Teachers record incidents of rule-breaking behaviors as they occur.	Students record rule-following behaviors on a variable-interval schedule.
System of reinforcement	Positive.	Positive.
Criterion for reinforcement	Teams must not exceed a set number of rule-breaking behaviors.	Groups or individuals achieve or exceed a set percentage of rule-following behaviors.
Delivery of reinforcement	Dependent on group performance.	Dependent on individual or group performance.
Feedback	Teacher provides feedback when rule-breaking behaviors occur.	Teacher provides feedback at intervals. Praise and encouragement are presented during the game to reinforce positive behaviors.

Determine the Behavior to Change and Any Collateral Behaviors That Might Be Affected

Let us suppose a dependent group contingency is established in which a class receives 10 minutes of extra free time, contingent on the improved academic performance of a student with developmental disabilities. Obviously, the teacher will need to collect data on the student's academic performance. However, data might also be collected on the number of positive interactions between the student and her classmates without disabilities within and outside the room. An additional benefit of using a group contingency might be the positive attention and encouragement that the student with developmental disabilities receives from her classmates.

Set Appropriate Performance Criteria

If a group contingency is used, the persons to whom the contingency is applied must have the prerequisite skills to perform the specified behaviors. Otherwise, they will not be able to meet the criterion and could be subject to ridicule or abuse (Stolz, 1978).

According to Hamblin, Hathaway, and Wodarski (1971), the criteria for a group contingency can be established by using the average-, high-, or low-performance levels of the groups as the standard. In an average-performance group contingency, the mean performance of the group is averaged, and reinforcement is contingent on the achievement of that mean score or a higher score. If the average score for a math exercise was 20 correct problems, a score of 20 or more would earn the reward. In a high-performance group contingency, the high score determines the level of performance needed to receive the reward. If the high score on a spelling test was 95%, only students achieving a score of 95% would receive the reward. In the low-performance group contingency, the low-performance score determines the reinforcer. If the low score on a social studies term paper was C, then students with a C or better would receive the reinforcer.

Hamblin et al. (1971) indicated that differential effects can be noted with these performance contingencies. Their data show that academically challenged students performed the worst under a high-performance contingency, whereas gifted students performed best under this contingency. The data of Hamblin and colleagues suggest that a group contingency can be effective in improving behavior, but it should be applied with the understanding that effectiveness may vary for different members of the group.

Combine with Other Procedures When Appropriate

According to LaRowe, Tucker, and McGuire (1980), a group contingency can be used alone or in combination with other procedures to systematically change performance. The LaRowe and colleagues study was designed to reduce excessive noise levels in an elementary school lunchroom; their data suggest that differential reinforcement of low rates of behavior (DRL) can be incorporated easily into a group contingency. In situations in which higher levels of group performance are desired, differential reinforcement of high rates of behavior (DRH) can be used. With either DRL or DRH, the use of a changing criterion design may facilitate the analysis of treatment effects.

Select the Most Appropriate Group Contingency

The selection of a specific group contingency should be based on the overall programmatic goals of the practitioner, the parents (if applicable), and the participants whenever possible. For instance, if the group contingency is designed to improve the behavior of one person or a small group of individuals, then perhaps a dependent group contingency should be employed. If the practitioner wants to differentially reinforce appropriate behavior, then an independent group contingency should be considered. But if the practitioner wants each individual within a group to perform at a certain level, then an interdependent group contingency should be chosen. Regardless of which group contingency is selected, the ethical issues discussed in Chapter 31 must be addressed.

Monitor Individual and Group Performance

When a group contingency is employed, practitioners must observe group and individual performance. Sometimes a group's performance improves, but members within the group do not improve, or at least do not improve as fast (Jowett Hirst et al., 2016). Some members of the group might even attempt to sabotage the group contingency, preventing the group from achieving the reinforcement. In these cases, individual contingencies should be arranged for the saboteurs that might combine the group contingency with other behavior reduction or behavior building procedures.

Future Applications of Group-Oriented Contingencies

Abundant evidence confirms that group-oriented contingencies have demonstrated effectiveness in classroom settings for individuals and groups, and across a range of instructional, therapeutic, and play-based settings (for a review, see Maggin, Johnson, Chafouleas, Ruberto, & Berggren, 2012). However, despite over 50 years of empirical demonstrations of their effectiveness, group contingencies, particularly the Good Behavior Game and its variations, have not been widely adopted, with an ever-wider range of populations yet to be explored (Vollmer, Joslyn, Rubow, Kronfli, & Donaldson, 2016).

Researchers might also turn their attention to evaluating group-oriented contingencies on larger issues facing society (e.g., global warming, recycling, reducing armed conflict, drug abuse). Some initial efforts using Internet-based group contingences have led to promising results with behaviors known to be systemic and addictive (Dallery, Raiff, & Grabinski, 2013; Dallery, Meredith, Jarvis, & Nuzzo, 2015).

CONTINGENCY CONTRACTING

A **contingency contract** (also called a *behavioral contract*) is a document that specifies a contingent relationship between the completion of a target behavior and access to, or delivery of, a specified reward. Contingency contracts stipulate how two or more people will behave toward each other. Such quid pro quo agreements make one person's behavior (e.g., preparing dinner) dependent on the other person's behavior (e.g., washing and putting away the dishes by a prescribed time the night before). Although verbal agreements may be considered contracts in the legal sense, they are not contingency contracts because the degree of specificity in designing, implementing, and evaluating a contingency contract far exceeds what is likely to occur in a verbal arrangement between parties. In addition, the act of signing a contingency contract and its prominent visibility during execution are integral parts of contingency contracts.

Contingency contracts have been used to improve academic performance (Wilkinson, 2003), weight control (Solanto, Jacobson, Heller, Golden, & Hertz, 1994), adherence to medical regimens (Miller & Stark, 1994), and athletic skills (Simek, O'Brien, & Figlerski, 1994). A compelling advantage of contingency contracts is their ability to be implemented alone or in packaged programs that incorporate two or more interventions concurrently (De Martini-Scully, Bray, & Kehle, 2000). In their meta-analysis of single-case research on behavior contracts, Bowman-Perrott, Burke, deMarin, Zhang, and Davis (2015) concluded that contingency contracts are beneficial for children and youth "regardless of grade level, gender, or disability status" (p. 247); are easily used by a wide range of practitioners; and can be combined with other behavior change interventions (e.g., De Martini-Scully et al., 2000; Navarro, Aguilar, Aguilar, Alcalde, & Marchena, 2007).

Contingency Contract Components

A contingency contract specifies the task and who is to perform it and the reward and who is to deliver it contingent on task completion. Figure 28.10 shows a contingency contract between a 10-year-old boy and his parents to help him get ready for school on time in the morning. This contract includes the two components required in all contingency contracts: task description and reward description—plus an optional task record.

Task

The task component consists of four parts. *Who* is the person who will perform the task and receive the reward—in this case, Lee. *What* is the task or behavior Lee must perform—getting ready for school. *When* identifies the time that the task must be completed—every school day. *How Well* is the most important part of the task component, and perhaps of the entire contract. It calls for the specifics of the task. Sometimes it is helpful to list a series of steps or subtasks so that the person can use the contract as a checklist of what must be done. Including a photo showing what the completed task should look like is helpful for some tasks (e.g., cleaning one's bedroom, setting the table for dinner). Any exceptions should be written in this part.

Reward

The reward side of a contract must be as complete and specific as the task side. Teachers and parents are usually good at specifying the task component of a contract; they know what they want their students or children to do. When it comes to the reward, however, specificity is lost and problems arise. Reward statements such as "Can watch some television" or "Will play catch when I get a chance" are not explicit, specific, or fair to the person completing the task.

On the reward side, *Who* is the person who will judge task completion and deliver the reward. With Lee's getting-ready-for-school contract, his mother is that person. *What* is the reward. *When* specifies the time that the reward can be received by the person earning it. With any contract it is crucial that the reward come *after* successful task completion. However, many rewards cannot be delivered immediately following task completion. In addition, some rewards have built-in, limited availability and can be delivered only at certain times (e.g., seeing the hometown baseball team play). Lee's contract specifies that his reward, if earned, will be received on Friday nights. *How Much* is the amount of reward that can be earned by completing the task. Any bonus contingencies should be included; for example, "By meeting her contract Monday through Friday, Ellie will earn bonus rewards on Saturday and Sunday."

CONTRACT

TASK

Who: Lee

What: Get ready for school

When: Each school day

How Well: Out of bed, dressed, cereal finished by 7:15. No more than 1 reminder from Mom or Dad.

REWARD

Who: Mom

What: Lee has a friend stay over

When: Friday night after perfect week

How Much: Friend can come home with Lee after school and stay overnight. Boys get pizza for late night treat.

Sign Here: Lee Behrle Date: Oct. 15, 2019

Sign Here: Audrey Behrle Date: Oct. 15, 2019

TASK RECORD

M	T	W	TH	F	M	T	W	TH	F	M	T	W	TH	F
★	★		★	★	★	★	★	★	★	★	★	★	★	★
		oops!	Great work, Lee					reward!!						reward

Figure 28.10 Example of a contingency contract.

Adapted from "*Sign Here: A Contracting Book for Children and Their Parents*" (2nd ed.) by J. C. Dardig and W. I. Heward, 2016. Copyright 1995 by J. C. Dardig and W. L. Heward. Used with Permission.

Task Record

Including a place on the contract to record task completion serves two purposes. First, recording task completion and reward delivery on the contract sets the occasion for all parties to review the contract regularly. Second, if a certain number of task completions are required to earn the reward, a checkmark, smiley face, or star can be placed on the task record each time the task is completed successfully. Marking the contract in this manner helps the person remain focused until the assignment is completed and the reward is earned. Lee's mother used the top row of boxes in the task record to record the days of the school week. In the middle row of boxes she placed gummed stars each day Lee met the conditions of his contract. In the bottom row, Lee's mother wrote comments about the progress of his contract.

Applications of Contingency Contracting

Contracting in the Classroom

Teachers have employed contracting to address specific discipline, performance, and academic challenges (De Martini-Scully et al., 2000; Flood & Wilder, 2002; Kehle, Bray, Theodore, Jenson, & Clark, 2000; Mruzek, Cohen, & Smith, 2007; Ruth, 1996). Newstrom and colleagues (1999), for instance, used a contingency contract with a middle school student with behavior disorders to improve the written mechanics associated with spelling and written language. After collecting baseline data on the percentage of capitalization and punctuation marks used correctly across spelling and sentence writing, a contingency contract was negotiated with and signed by the student that specified that improved performance would yield free time on the classroom computer. The student was reminded of the terms of the contract before each language arts class that included spelling worksheets and journal writing (i.e., sentences).

Figure 28.11 shows the results of the contingency contracting intervention. When baseline was in effect for spelling and written sentences respectively, mean scores for both variables were in the 20% correct range. When contracting was initiated, the student's performance for spelling and written sentences increased immediately to an average of approximately 84% correct. The authors also reported positive anecdotal evidence related to spelling and written language from other teachers with whom this student interacted.

Wilkinson (2003) used a contingency contract to reduce the disruptive behavior of a first-grade student. Disruptive behaviors included being off task, refusals to comply with work assignments and instructions, fighting with peers, and temper tantrums. A behavioral consultation effort was launched with the classroom teacher that included problem identification, analysis, intervention, and evaluation. Contingency contracting consisted of the student earning preferred rewards and social praise from the teacher for three behaviors: increased time on task, appropriate interactions with other children, and compliance with teacher requests. Observations of her behavior over 13 sessions of baseline and contingency contracting showed a decrease in the percentage of intervals with disruptive behavior when the contingency contract was in effect. Wilkinson reported

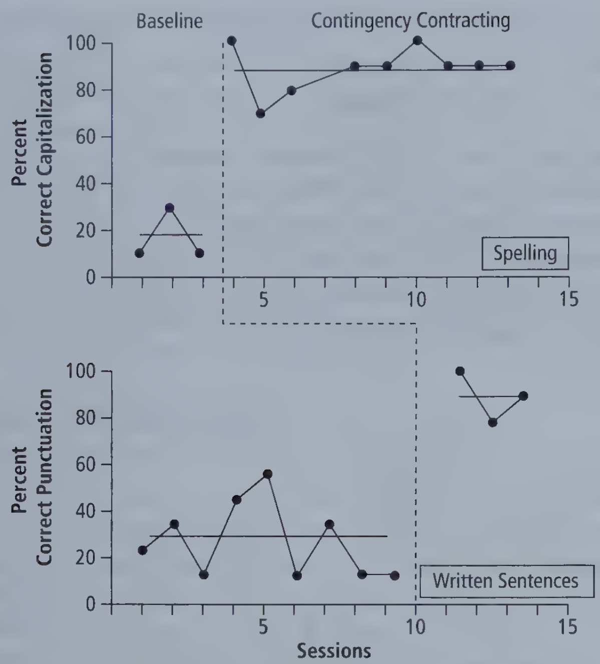


Figure 28.11 Percentage of capitalization and punctuation marks used correctly on spelling worksheets and journal writing during baseline and contingency contracting.

From "The Effects of Contingency Contracting to Improve the Mechanics of Written Language with a Middle School Student with Behavior Disorders," by J. Newstrom, T. F. McLaughlin, and W. J. Sweeney, 1999, *Child & Family Behavior Therapy*, *21*(1), p. 44. Copyright 1999 by The Haworth Press, Inc. Reprinted by permission.

that the student's disruptive behavior decreased substantially and remained low during a 4-week follow-up period.

Ruth (1996) conducted a 5-year longitudinal study that blended contingency contracting with goal setting with students who had emotional and behavioral problems. After students negotiated contracts with their teachers, a goal-setting component was added that included statements about their daily and weekly goals and the criterion levels for success. Results for the 37 out of 43 students who finished the program after 5 years showed that 75% of daily goals, 72% of weekly goals, and 86% of total goals were reached. Ruth summarized the beneficial effects of combining strategies: "When [goal-setting] methods are incorporated into a contract, the motivational aspects of behavior contracting and goal setting may combine to produce maximum effort and success" (p. 156).

Contracting in the Home

Miller and Kelley (1994) combined contingency contracting and goal setting to improve the homework performance of four preadolescent students with histories of poor homework completion and who were at risk for other academic problems (e.g., procrastination, being off task, errors with submitted work). During baseline, parents recorded their children's homework time on task, type and accuracy of problem completion, and number of problems completed correctly. Subsequently, parents and children entered into a goal setting and contingency contracting phase that was preceded by parent training on how to set and negotiate goals and write contracts. Each night, parents and children established their respective goals and negotiated a compromise goal based on that interaction. Each week, renegotiations occurred for tasks, rewards, and sanctions should the contract not be met. A recording sheet was used to measure progress.

Figure 28.12 shows the results of the study. When goal setting was combined with contingency contracting, accuracy performance increased for all students. Miller and Kelley's findings reaffirm the notion that contingency contracting can be combined successfully with other strategies to produce functional outcomes.

Clinical Applications of Contracting

Flood and Wilder (2002) combined contingency contracting with functional communication training to reduce the off-task behavior of an elementary-aged student diagnosed with attention-deficit hyperactivity disorder (ADHD) who had been referred to a clinic-based program because his off-task behavior had reached alarming levels. Antecedent assessment, functional communication training, and contingency contracting were conducted in a therapy room located in the clinical facility. Specifically, antecedent assessments were conducted to determine the level of off-task behavior when the difficulty of academic tasks varied from easy to difficult and therapist attention varied from low to high. A preference assessment was also conducted. Using discrete trial training, the student was taught to raise his hand for assistance with tasks (e.g., "Can you help me with this problem?"). The therapist sat nearby and responded to appropriate requests for assistance and ignored other vocalizations. Once requesting assistance was mastered, the contingency contract was established whereby the student could earn preferred items, identified through the assessment, contingent on accurate task completion. The results showed that during baseline, off-task performance was high for math division and word problems. When the intervention was introduced, an immediate reduction in off-task behavior was noted for division and word problems. Also, the student's accuracy in completing the division and word problems improved as well. Whereas during baseline conditions he solved correctly 5% and 33% of the division and word problems, respectively, during intervention he solved 24% and 92% of the problems correctly.

Using Contracting to Teach Self-Management Skills to Children

Ideally, contingency contracting involves the active participation of the child throughout the development, implementation, and evaluation of the contract. For many children, contracting is a first experience in identifying specific ways they would like to act and then arranging certain aspects of their environment to set the occasion for and reward those acts. If more of the decision making for all parts of the contracting process is turned over to children gradually and systematically, they can become skilled at self-contracting. A **self-contract** is a contingency contract that a person makes with herself, incorporating a self-selected task and reward as well as personal monitoring of task completion and self-delivery of the reward. Self-contracting skills can be achieved by a multistep process of having an adult prescribe virtually all of the elements of the task and reward and gradually shifting the design of the elements to the child.

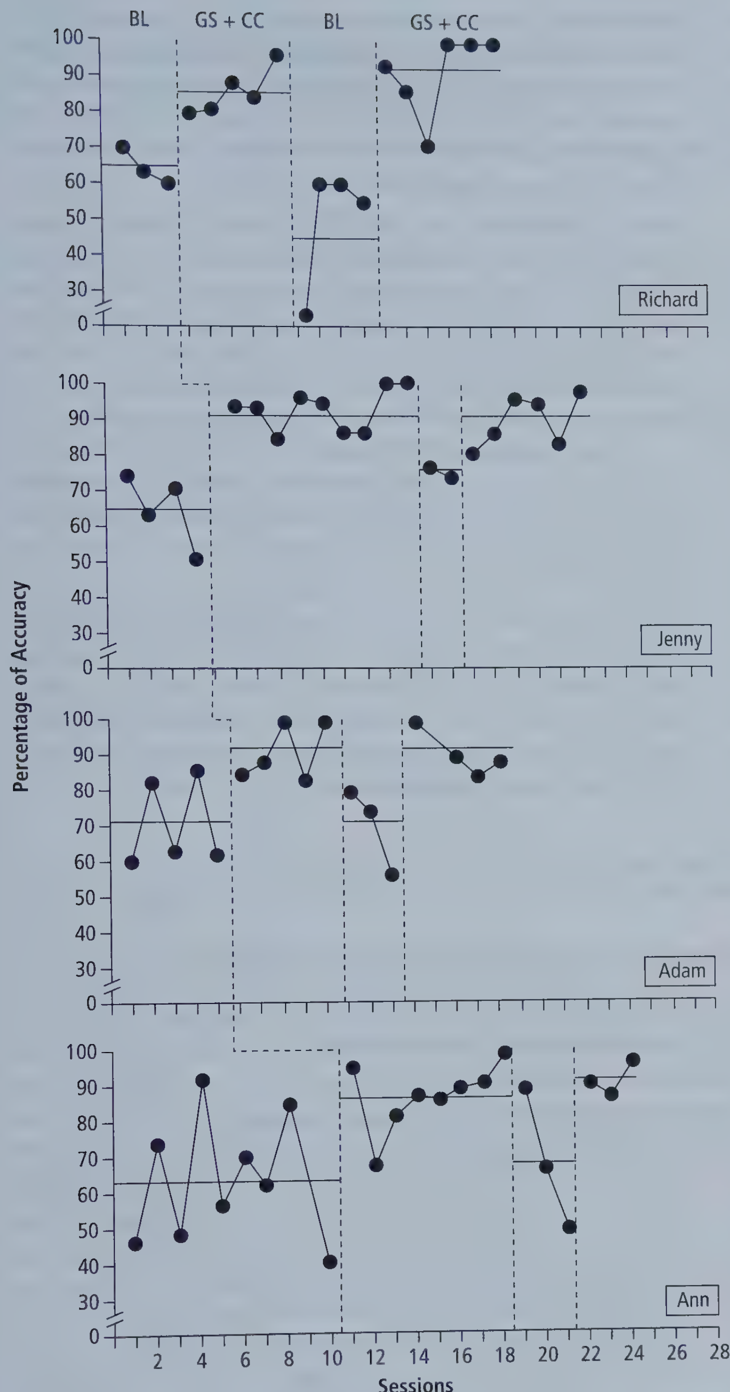


Figure 28.12 Percentage of homework problems completed accurately during baseline and a treatment condition consisting of goal setting and contingency contracting. Sessions correspond to sequential school days (i.e., Mondays through Thursdays) on which subjects were assigned homework. Data were not collected on days on which homework was not assigned.

From "The Use of Goal Setting and Contingency Contracting for Improving Children's Homework Performance" by D. L. Miller & M. L. Kelley, 1994, *Journal of Applied Behavior Analysis, 27*, p. 80. Reproduced with permission of John Wiley & Sons,Inc.

How Do Contracts Work?

At first glance the principle of behavior behind contingency contracts seems deceptively simple: A behavior is followed by a contingent reward—surely a case of positive reinforcement. Yet, in most contracts the reward, although contingent, is too delayed to reinforce task completion directly; and many successful contracts specify a reward that would not function as a reinforcer for the task even if it was presented immediately after the task completion. Further, behavioral contracting is not a single procedure with a single associated behavior and a single reinforcer. Contracting is more accurately conceptualized as an intervention package that combines several behavior principles and procedures.

So, how do contracts work? Several principles, procedures, and factors are likely to apply. Certainly, reinforcement is involved, but not in as simple or direct a fashion as it might seem at first. *Rule-governed behavior* is probably involved (Malott & Shane, 2014; Skinner, 1969; see Chapter 11). A contract describes a rule: A specified behavior will be followed by a specified (and reasonably immediate) consequence. The contract itself serves as a response prompt to perform the target behavior and enables the effective use of a consequence (e.g., going to the movies Saturday night), too delayed to reinforce certain behaviors (e.g., practicing the trumpet on Tuesday). Delayed consequences can help exert control over behaviors performed hours and even days before if they are associated with and linked

by verbal behavior to the rule (e.g., "I've just finished my trumpet practice—that's another checkmark toward the movies on Saturday") or to interim token reinforcers (e.g., the checkmark on the contract after practicing). The physical visibility of the contract may also function as a response prompt for escaping "guilt" (Malott & Shane, 2014).

Developing Contingency Contracts

Although teachers, therapists, or parents can unilaterally determine a contract for a child or client, contracting is usually more effective when all of the parties involved play an active role in developing the contract. Contract development involves the specification of tasks and rewards in a fashion agreeable and beneficial to each party. Dardig and Heward (2016) described a five-step procedure for identifying tasks and rewards that can be used by teachers and families.[2]

Step 1: Hold a meeting. To get the entire group (family or class) involved in the contracting process, a meeting should be held. At this meeting, members can discuss how contracts work, how they can help the group cooperate and get along better, and how contracts can help individuals meet personal goals. Parents or teachers should emphasize that they will participate in all of the steps leading up to and including implementation of contracts. It is important that children view contracting as a behavior-exchange process shared by all members of the group, not as something adults impose on them. The field-tested list-making procedures described in the following steps provide a simple and logical framework for the selection of tasks and rewards for family and classroom contracts. Most groups can complete the procedure within 1 to 2 hours.

Step 2: Complete List A. Each member completes three lists prior to the actual writing of the contract. List A (see Figure 28.13) is designed to help each member identify not only those tasks he can perform within the context of a contract, but also those tasks he already does to help the group. In this way, positive attention can be focused on appropriate behaviors that individual members are currently completing satisfactorily.

Each member should be given a copy of List A. Everyone should be careful to describe all tasks as specifically as possible. Then the completed lists can be put aside, and the group can proceed to the next step. If a member is unable to write, that person's list can be completed orally.

Step 3: Complete List B. List B (see Figure 28.14) is designed to help group members identify possible contract tasks for other group members and helpful behaviors currently being completed by those persons. List B can also identify areas where disagreement exists between group members as to whether certain tasks are actually being completed properly and regularly.

Figure 28.13 A form for self-identification of possible tasks for contingency contracts.

List A Name: Charlotte

THINGS I DO TO HELP MY FAMILY	OTHER WAYS I COULD HELP MY FAMILY AND MYSELF
1. Entertain us with skits	1. Be on time for dinner
2. Clean up my bedroom	2. Turn off the lights when I leave a room
3. Practice my violin	3. Follow morning routine
4. Wash dishes	4. Hang up my coat when I get home from school
5. Help Dad fold the laundry	5. Listen to Mom and Dad the first time
6.	6.
7.	7.

From *Sign Here: A Contracting Book for Children and Their Parents* (2nd ed., p. 111) by J. C. Dardig and W. L. Heward, 1981. Bridgewater, NJ: Fournies and Associates. Copyright 1981 by Fournies and Associates. Reprinted by permission.

Figure 28.14 A form for identifying potential contracting tasks for other family members.

List B Name: Bobby

THINGS BOBBY DOES TO HELP THE FAMILY	OTHER WAYS BOBBY COULD HELP THE FAMILY
1. Vacuums when asked	1. Put his dirty clothes in hamper
2. Makes his bed	2. Do homework without being asked
3. Reads stories to little sister	3. Make his own sandwiches for his school lunch
4. Empties trash	4. Clean and sponge off table after dinner
5. Rakes leaves	5.
6.	6.
7.	7.

From *Sign Here: A Contracting Book for Children and Their Parents* (2nd ed., p. 111) by J. C. Dardig and W. L. Heward, 1981. Bridgewater, NJ: Fournies and Associates. Copyright 1981 by Fournies and Associates. Reprinted by permission.

Figure 28.15 A form for self-identification of possible rewards for contingency contracts.

List C **Name:** Charlotte

MY FAVORITE THINGS, ACTIVITIES, AND SPECIAL TREATS

1. Listening to music
2. Movies
3. New book
4. Sleepover with friend
5. Swimming
6. Special packed lunches for school
7. Ice cream sundaes
8. Toys for Kiley (pet dog)
9. Picnics
10. New shoes
11. ______
12. ______
13. ______
14. ______
15. ______

From *Sign Here: A Contracting Book for Children and Their Parents* (2nd ed., p. 111) by J. C. Dardig and W. L. Heward, 1981. Bridgewater, NJ: Fournies and Associates. Copyright 1981 by Fournies and Associates. Reprinted by permission.

Each member should be given a copy of List B and asked to write his or her name in all three blanks at the top. These lists can then be passed around the table so that everyone has a chance to write at least one behavior on each side of everyone else's list. Everyone writes on every List B except his or her own, and each person should be required to write at least one positive behavior on everyone else's List B. After completion, these lists should be set aside before moving to the next step.

Step 4: Complete List C. List C (see Figure 28.15) is simply a sheet with numbered lines on which each group member identifies potential rewards he would like to earn by completing contracted tasks. Participants should list not only everyday favorite things and activities but also special items and activities they may have wanted for a long time. It is all right if two or more people indicate the same reward. After List C is completed, each person should collect his two other lists and read them carefully, talking over any misunderstood items.

Step 5: Write and Sign Contracts. The final step begins with choosing a task for each person's first contract. Discussion should move around the group, with members helping each other decide which task is the most important to start doing first. Everyone should write *who* is going to perform the task, exactly *what* the task is, *how well* and *when* it has to be done, and any possible exceptions. Each person should also look at List C and choose a reward that is neither excessive nor insignificant but is fair for the selected task. Each member should write *who* will control the reward, *what* the reward is, *when* it is to be given, and *how much* is to be given. Everyone in the group should write and sign one contract during the first meeting.

Contracting with Nonreaders

Reading ability by the participant is not a prerequisite for contingency contracting; however, the individual must be able to come under the control of the visual or oral statements (rules) of the contract. Contracting with nonreaders involves three types of clients: (a) preschoolers with good verbal skills, (b) school-aged children with limited reading skills, and (c) adults with adequate language and conceptual skills but who lack reading and writing skills. Contracts for nonreaders specify tasks and rewards with nonprint elements such as symbols, pictures, photos, and audio recordings chosen to suit the individual skills and preferences of children and adults in all three nonreader groups (see Figure 28.16).

Figure 28.16 A contingency contract for a nonreader.

Contracting with Nonsigners

Although some children are eager, and most at least willing, to try a contract, some want nothing to do with the whole idea. Using contingency contracting in a collaborative approach (Lassman, Jolivette, & Wehby, 1999) may reduce the likelihood of noncompliance, but a parent or practitioner should never try to force the child to sign a contract. Coercion is contrary to the whole idea of behavioral contracting. Instead of giving up, however, Dardig and Heward (2016) suggest the following strategies for working with a child who refuses to enter into a contingency contract:

- *Shaping and ignoring.* Remember the nonsigner's anti-contracting behavior is just that—*behavior*. Perhaps the worst thing parents can do when their child says "Contracts are stupid!" or "You'll never get me to sign a contract" is to argue or try to convince the child that contracts are "cool." Better instead to pay attention to the child when he makes any remotely positive or even neutral comments about contracting, such as, "I wonder if anyone has ever made a contract with his parents for _________."
- *Modeling.* In families with other children, perhaps the nonsigner's sibling will try contracting. In a two-parent family, the parents can do contracts with one another; a single parent might develop a contract with a friend as a way to demonstrate the idea. A nonsigner who observes his sibling(s) and/or parent(s) successfully using a contract may then be willing to give it a try.
- *Parent contract.* Past experience may lead children to be wary of what their parents have in store for them and to always believe they get the short end of the stick. The parent might say something like, "Okay, if you don't think contracts are fair (or whatever the nonsigner doesn't like about contracts), let's make the first contract be for my (our) behavior and you control the reward." After controlling the reward for a contract specifying a task his parent(s) must do, the nonsigner may learn that contracts can work, and may be willing to enter into a contract for his behavior.
- *Self-contract.* Perhaps the nonsigner does not trust that the contract will be honest, that he will complete the task but the reward will not be delivered. Allowing the child to control both the task and the reward components of a contract is one tactic for dealing with this suspicion. Will the child cheat if given such freedom? Perhaps the child will, but most will not. Some research shows that self-determined standards of task and reward result in better performance by children than do criteria set by adults (Dickerson & Creedon, 1981; Olympia, Sheridan, Jenson, & Andrews, 1994).

However, even after the best of positive approaches, the reality is that some nonsigners may not agree to participate in a contingency contract. In such cases, another behavior change strategy would be a better alternative for dealing with the target behavior.

Guidelines for Implementing and Evaluating Contracts

In determining whether contingency contracting is an appropriate intervention for a given problem, the practitioner should consider the nature of the desired behavior change, the verbal and conceptual skills of the participant, the individual's relationship with the person(s) with whom the contract will be made, and the available resources. The target behavior to be changed by a contingency contract should already be in the person's repertoire and under proper stimulus control in the settings where the response is desired. If the behavior is not in the individual's repertoire, other behavior-building techniques should be attempted (e.g., modeling, shaping, chaining). Contracting is most effective with behaviors that produce permanent products (e.g., completed homework assignment, cleaned bedroom) or that occur in the presence of the person who is to deliver the reward (e.g., the teacher or parent).

Numerous lists of rules and guidelines for effective contingency contracting have been published (e.g., Dardig & Heward, 2016; Downing, 1990; Homme, Csanyi, Gonzales, & Rechs, 1970). Table 28.3 provides a list of the frequently cited guidelines and rules.

The simplest way to evaluate a contract is to record the occurrence of task completion. Including a task record on the contract helps make evaluation a natural by-product of the contracting process. By comparing data from the task record with a precontract baseline of task completion, an objective determination can be made as to whether improvement has occurred. A good outcome is one that results in the specified task being completed more often than it was before the contract.

Sometimes the outcome data indicate that the task is being completed more often and more consistently than it was before the contract, but the parties involved are unsatisfied. In such cases, either the original problem or goal that prompted the development of the contract is not being attained, or one or more of the participants are unhappy with how the contract is being carried out. The first possibility results from selecting the wrong behavior for the task part of the contract. For example, suppose ninth-grader Lenwood wants to improve on the D's and F's he has been getting on his algebra assignments at school and writes a contract with his parents, specifying the task as "studying algebra for 1 hour each school night." After several weeks, Lenwood has studied algebra for the required 1 hour on all but two school nights, but his poor grades remain unchanged. Has Lenwood's contract worked? The correct answer is both yes and no. Lenwood's contract was successful in that he consistently completed the specified task—1 hour of study each school night. However, in terms of his original objective—earning better algebra grades—the contract was a failure. Lenwood's contract helped him change the behavior he specified, but he specified the wrong behavior. Studying for 1 hour, for Lenwood at least, was not directly related to his goal. By changing his contract to require that he solve correctly 10 algebra equations each night (the behavior required to get good grades on algebra tests), his goal of obtaining better grades may become a reality.

It is also important to consider the participant's reactions to the contract. A contract that produces desired change in the specified target behavior but causes other maladaptive or emotional responses may be an unacceptable solution. Having the client participate in the development of the contract and jointly conducting regular progress checks help to avoid this situation.

TABLE 28.3 Contingency Contracting Rules and Guidelines

Rules and Guidelines	Comments
Write a fair contract.	There must be a fair relationship between the difficulty of the task and the amount of the reward. The goal is to achieve a win–win situation for both parties, not for one party to gain an advantage over the other.
Write a clear contract.	In many instances, a contract's greatest advantage is that it specifies each person's expectations. When a teacher or parent's expectations are explicit, performance is more likely to improve. Contingency contracts must say what they mean and mean what they say.
Write an honest contract.	An honest contract exists if the reward is delivered at the time and in the amount specified when the task is completed as agreed. In an honest contract, the reward is *not* delivered if the task has not been completed as specified.
Build in several layers of rewards.	Contracts can include bonus rewards for beating the best daily, weekly, or monthly performance.
Consider a response cost contingency.	Occasionally, it may be necessary to incorporate a "fine"—the removal of rewards—if the agreed-upon task is not completed.
Post the contract in a visible place.	Public posting allows all parties to see progress toward achieving the goals of the contract.
Renegotiate and change a contract when either party is consistently unhappy with it.	Contracting is designed to be a positive experience for all parties, not a tedious endurance contest to determine survivors. If the contract is not working, reconsider the task, the reward components, or both.
Terminate a contingency contract.	A contingency contract is a means to an end, not the end product. Once independent and proficient performance is achieved, the contract can be terminated. Further, a contract can and should be terminated when one party or both parties consistently fail to meet the contract's terms.

SUMMARY

Token Economy

1. A token economy, also called a token reinforcement system, is a behavior change system consisting of three major components: (a) a specified list of target behaviors to be reinforced; (b) tokens or points that participants receive for emitting the target behaviors; and (c) a menu of items or activities, privileges, and backup reinforcers from which participants choose and obtain by exchanging tokens they have earned.

2. Tokens are objects, symbols, marks, or chips that function as generalized conditioned reinforcers because they have been paired with a wide variety of backup reinforcers.

3. As participants move "up" from one level to the next, they have access to more privileges and are expected to demonstrate more independence. A level system is a type of token economy in which participants move up or down a hierarchy of levels contingent on meeting specific performance criteria with respect to the target behaviors.

4. There are six basic steps in designing a token economy: (1) identifying the target behaviors and rules; (2) selecting the tokens that will serve as a medium of exchange; (3) selecting a menu of backup reinforcers; (4) establishing a token exchange ratio; (5) writing procedures to specify when and how tokens will be dispensed and exchanged and what will happen if the requirements to earn a token are not met; and (6) field-testing the system before full-scale implementation.

5. When establishing a token economy, decisions must be made regarding how to begin, conduct, maintain, remove, and evaluate the system.

Group Contingencies

6. In a group contingency a common consequence (usually, but not necessarily, a reward intended to function as reinforcement) is contingent on the behavior of one member of the group, the behavior of part of the group, or the behavior of everyone in the group.

7. Group contingencies can be classified as independent, dependent, and interdependent, and they offer the behavior analyst multiple advantages for improving or reducing behaviors.

8. Six guidelines can assist the practitioner in implementing a group contingency: (a) choose a powerful reward, (b) determine the behavior to change and any collateral behaviors that might be affected, (c) set appropriate performance criteria, (d) combine with other procedures when appropriate, (e) select the most appropriate group contingency, and (f) monitor individual and group performance.

9. Future applications of group contingencies might explore their effectiveness with larger societal issues (e.g., global warming), or chronic, systemic, difficult-to-treat individual challenges (e.g., smoking cessation, addictive behaviors).

Contingency Contracting

10. A contingency contract is a document that specifies a contingent relation between the completion of a target behavior and access to, or delivery of, a specified reward.

11. Two components required in all contingency contracts are a description of the task and a description of the reward. Under *task*, who, what, when, and how well should be specified. Under *reward*, who, what, when, and how much should be specified. An optional task record provides a place for recording the progress of the contract and providing interim rewards.

12. A self-contract is a contingency contract that an individual makes with herself, incorporating a self-selected task and reward as well as self-monitoring of task completion and self-delivery of the reward.

13. Contracts for nonreaders specify tasks and rewards with nonprint elements such as symbols, pictures, photos, and audio recordings.

KEY TERMS

backup reinforcer	group contingency	level system
behavioral contract	hero procedure	self-contract
contingency contract	independent group contingency	token
dependent group contingency	interdependent group contingency	token economy

MULTIPLE-CHOICE QUESTIONS

1. A description of the task should include which of the following when writing a contingency contract?
   a. A description of the reward
   b. A description of when the reward will be delivered
   c. A method of tracking progress
   d. A description of when the task should be performed
   Hint: (See "Contingency Contracting")

2. A contingency contract:
   a. Is a verbal agreement among two individuals
   b. Specifies how two people will behave toward each other
   c. Often causes scapegoating among learners and peers
   d. Depends largely on the use of punishment for its effectiveness
   Hint: (See "Contingency Contracting")

3. Reinforcers in a token economy might include:
   a. Naturally occurring activities such as classroom jobs, games, etc.
   b. Privileges
   c. Tangible items, such as candy, pencils, etc.
   d. All of these
   Hint: (See "Token Economy")

4. The primary reason one might consider using a level system with a token system is that:
   a. Doing so can foster self-management
   b. Doing so makes it easier to withdraw the token economy
   c. Level systems are more preferred for all learners
   d. Level systems are easier to manage
   Hint: (See "Token Economy")

5. Token economies often need to be:
   a. Complex
   b. Individualized
   c. Maintained forever
   d. All of these
   Hint: (See "Implementing a Token Economy")

6. When selecting backup reinforcers:
   a. Consider ethical and legal issues
   b. It is a good idea to use basic need items (e.g., lunch) as a reinforcer
   c. Keep the pool of backup reinforcers small
   d. Never use edibles
   Hint: (See "Implementing a Token Economy")

7. A teacher says to his class: "Melissa is working on remembering to follow instructions when I give them. Melissa is going to keep track of whether or not she follows my instructions. If she follows 90% of my instructions today (or better), our whole class will watch a movie at the end of the day." What type of group contingency is this?
   a. Independent
   b. Dependent
   c. Interdependent
   d. None of these
   Hint: (See "Group Contingencies")

8. Group contingencies can:
   a. Lighten the workload for the practitioner
   b. Save time
   c. Create problems with learners
   d. All of these
   Hint: (See "Group Contingencies")

9. How could a learner sabotage a group contingency?

a. By purposefully not meeting the performance criterion so that no one gets the reinforcer

b. By not showing up for school on a given day

c. By urging peers to participate in the program

d. By congratulating peers who are meeting the criterion

Hint: (See "Group Contingencies")

10. An interdependent group contingency is most useful if:

a. One wants to change the behavior of one person or a small group of individuals

b. One wants to differentially reinforce individuals for appropriate behavior

c. One wants to change everyone's performance

d. All of these

Hint: (See "Group Contingencies")

ESSAY-TYPE QUESTIONS

1. When describing the reward in a contingency contract, what are the four important things to include in the description?

Hint: (See "Contingency Contracting")

2. You are a consultant working with a family that wants to write a contingency contract with their children to have better behavior in the home. Describe the process you would ask the family to go through to develop the contract.

Hint: (See "Developing Contingency Contracts")

3. What are the main components of a token economy?

Hint: (See "Token Economy")

4. If you were going to implement a token economy, what would you use as a token? Why would you select this as a token? Make sure in your answer you discuss the important considerations one must make when selecting the token.

Hint: (See "Token Economy")

5. Identify and define the different types of group contingencies.

Hint: (See "Group Contingencies")

6. What are the six guidelines for implementing a group contingency?

Hint: (See "Guidelines for Implementing a Group Contingency")

NOTES

1. A total of 1500 steps was added, as researchers determined that this number of steps represented 15 to 20 minutes of at least moderate exercise per day.
2. Research and development of the *Sign Here* contracting method described here was informed by three doctoral dissertations at The Ohio State University. Kabler (1976) used *Sign Here* to teach contracting skills to three classes of fourth-grade students. Norman (1977) found that eight of nine parents of children exhibiting behavior problems at home wrote and implemented contracts that reduced disruptions in the home. Shrewsberry (1977) used *Sign Here* to teach contracting to parents from 59 families, who wrote and implemented 154 contracts during a 6-week period. The parents judged 138 (90%) of those contracts as successful in increasing their children's completion of the specified tasks.

CHAPTER 29

Self-Management

LEARNING OBJECTIVES

- Define self-management, its advantages and benefits, and specific types of self-management strategies.
- Define self-monitoring, list and discuss suggested guidelines or procedures for its use, and discuss various derivatives of self-monitoring.
- Compare and contrast different types of self-administered consequences.
- Define and discuss four additional types of self-management tactics.
- List and discuss various suggestions for conducting an effective self-management program.

Raylene often failed to do the things she needed and wanted to do. Her days were just so busy! But Raylene is beginning to get a handle on her hectic life. This morning a sticky-note on her closet door reminded her to wear her gray suit for her luncheon meeting. A note on the refrigerator prompted her to take her completed sales report to work. When Raylene got into her car—sales report in hand and looking sharp in her gray suit—and sat on the library book she'd placed on the driver's seat the night before, chances were good that she would return the book that day and avoid another overdue fine.

Daryl collected the last data point for his master's thesis nearly a year ago. It's a solid study on a topic he enjoys and believes important, but Daryl is struggling mightily to write his thesis. Daryl knows if he would sit down and write for an hour or two a day his thesis would get written. But the scope and difficulty of the task are daunting. Daryl wishes his ability to sit and write were half as well developed as his propensity for distraction.

Raylene recently discovered self-management—behaving in ways that alter the occurrence of other behaviors she wants to control—and feels on top of the world. Daryl sorely needs some self-management. This chapter defines self-management, identifies uses for self-management and the benefits of teaching and learning self-management skills, describes a variety of self-management strategies and tactics, and offers guidelines for designing and implementing successful self-management programs. We begin with a discussion of the role of the self as a controller of behavior.

THE "SELF" AS BEHAVIOR CONTROLLER

A fundamental precept of radical behaviorism is that the causes of behavior are found in the environment. Throughout the evolution of the human species, causal variables selected by the contingencies of survival have been transmitted through genetic endowment. Other causes of behavior are found in the contingencies of reinforcement that operate during the lifetime of an individual. What role, then, if any, is left for the *self* to play?

Locus of Control: Internal or External Causes of Behavior

Proximate causes of some behaviors are apparent by observing events as they unfold. A mother picks up and cuddles her crying infant, and the baby stops crying. A highway worker leaps off the road when a fast-approaching car nears. An angler casts his lure toward the spot where his previous cast produced a strike. A behavior analyst would likely suggest the involvement of escape, avoidance, and positive reinforcement contingencies, respectively, for these events. Although others might offer mentalistic explanations for why the person in each scenario responded as he or she did (e.g., the infant's crying triggered the mother's nurturing instinct), most people would identify the same antecedent events—a baby cried, a speeding car approached, a fish struck the lure—as functional variables in the three scenarios. An analysis of the three episodes would reveal almost certainly that the immediately preceding events had functional roles: the crying baby and speeding car as motivating operations that evoked escape and avoidance responses; the biting fish as a reinforcer that led to repeating the cast that produced it.

But much human behavior does not immediately follow such obviously related antecedent events. Nonetheless, we humans have a long history of assigning causal status to events that immediately precede behavior. As Skinner (1974) noted, "We tend to say, often rashly, that if one thing follows another, it was probably caused by it—following the ancient precept of *post hoc, ergo propter hoc* (after this, therefore because of this)" (p. 7). When causal variables are not readily apparent in the immediate environment, the tendency to point to internal causes of behavior is particularly strong. As Skinner explained:

> The person with whom we are most familiar is our self; many of the things we observe just before we behave occur

Figure 29.1 Examples of four target behaviors for self-management and the self-management behavior emitted to change each example.

Target Behavior	Self-Management Behavior
• Save money instead of spending it on frivolous items.	• Enroll in a payroll deduction plan.
• Take garbage and recycling cans from garage to the curb on Thursday nights for pickup early Friday morning.	• After backing the car out of the garage for work Thursday morning, pull the trash and recycling cans onto the space in the garage where you will park the car in the evening.
• Ride exercise bike for 30 minutes each evening.	• Make a chart of the minutes you ride, and show it to a coworker each morning.
• Write a 20-page term paper.	• (1) Outline paper and divide it into five 4-page parts; (2) specify a due date for each part; (3) give roommate five $10 checks made out to a despised organization; (4) on each due date show a completed part of the paper to roommate and get one $10 check back.

> within our body, and it is easy to take them as the causes of our behavior. . . . Feelings occur at just the right time to serve as causes of behavior, and they have been cited as such for centuries. We assume that other people feel as we feel when they behave as we behave. (pp. 7, 8)

Why does one college student maintain a regular study schedule from the first week of the semester while her roommates party night after night? When a group of people joins a weight loss or a smoking cessation program in which each member is exposed to the same intervention, why do some people meet their self-determined goals but others do not? Why does a high school basketball player with limited physical abilities consistently outperform his more athletically gifted teammates? The hardworking student is said to have more willpower than her less studious roommates; the group members who lost weight or stopped smoking are thought to possess more desire than their peers who failed to meet their goals or dropped out; the athlete's superior play is considered a result of his exceptional drive. Although some psychological theories grant causal status to hypothetical constructs such as *willpower, desire*, and *drive*, these explanatory fictions lead to circular reasoning, bring us no closer to understanding the behaviors they claim to explain, and thus have no role in a behavior analysis of self-control.[1]

Skinner's Two-response Conceptualization of Self-Control

Skinner was the first to apply the philosophy and theory of radical behaviorism to behavior typically considered controlled by the self. In his classic textbook, *Science and Human Behavior*, Skinner (1953) devoted a chapter to self-control.

> When a man controls himself, chooses a course of action, thinks out the solution to a problem, or strives toward an increase in self-knowledge, he is *behaving*. He controls himself precisely as he would control the behavior of anyone else—through the manipulation of variables of which behavior is a function. His behavior in so doing is a proper object of analysis, and eventually it must be accounted for with variables lying outside the individual himself. (pp. 228–229)

Skinner (1953) conceptualized **self-control [Skinner's analysis]** as a two-response phenomenon:

> One response, the *controlling response*, affects variables in such a way as to change the probability of the other, the *controlled response*. The controlling response may manipulate any of the variables of which the controlled response is a function; hence there are a good many different forms of self-control. (p. 231)

Skinner (1953) described a wide variety of self-control techniques, including use of physical restraint (e.g., clapping a hand over one's mouth to prevent yawning at an embarrassing moment), changing the antecedent stimulus (e.g., putting a box of candy out of sight to reduce overeating), and "doing something else" (e.g., talking about another topic to avoid talking about a particular topic), to name a few. Since Skinner's initial list of techniques, a variety of taxonomies of self-control tactics have been described (e.g., Kazdin, 2013; Watson & Tharp, 2014). All self-control—or self-management—tactics can be operationalized in terms of two behaviors: (a) the target behavior a person wants to change (Skinner's controlled response) and (b) the self-management behavior (Skinner's controlling response) emitted to change the target behavior. Consider the examples in Figure 29.1.

SELF-MANAGEMENT DEFINED

There is nothing mystical about self-management.[2] Self-management is behavior a person emits to influence another behavior. But many of the responses a person emits each day affect other behaviors. Putting toothpaste on a toothbrush makes it likely that tooth brushing will occur soon. But we do not consider squeezing toothpaste on a brush an act of "self-management" just because it evokes tooth brushing. What gives some responses the special status of self-management? How can we distinguish self-management from other behavior?

Numerous definitions of self-control or self-management have been proposed, many of which are similar to one offered by Thoresen and Mahoney (1974). They suggested that self-control occurs when, in the "relative absence" of immediate external

controls, a person emits a response designed to control another behavior. For example, a man shows self-control if, when home alone and otherwise free to do whatever he wants, he forgoes his usual peanuts and beer and rides his exercise bike for 30 minutes. According to definitions such as Thoresen and Mahoney's, the man would not be credited with self-control if his wife were there reminding him not to overeat, praising him for riding the exercise bike, and marking his time and mileage ridden on a chart.

What, however, if the man had asked his wife to prompt, praise, and chart his exercise? Would his having done so constitute a form of self-control? One problem with conceptualizing self-control as occurring only in the absence of "external control" is that it excludes those situations in which a person designs contingencies that entail external control for a behavior he desires to change. Another problem with the "relative absence of external controls" concept of self-control is that it creates a false distinction between internal and external controlling variables when, in fact, all causal variables for behavior ultimately reside in the environment.

Definitions of self-control such as Kazdin's (2013), "those behaviors that an individual deliberately undertakes to achieve self-selected outcomes by manipulating antecedent and consequence events" (p. 627), are more functional for applied behavior analysts. With this definition, self-control occurs whenever a person purposely emits behavior that changes the environment to modify another behavior. Self-control is considered *purposeful* in the sense that a person labels (or *tacts*) her responses as designed to attain a specified result (e.g., to reduce the number of cigarettes smoked each day).

We define **self-management** as the personal application of behavior change tactics that produces a desired improvement in behavior. This intentionally broad definition of self-management encompasses one-time self-management events, such as Raylene's taping a note to her closet door to remind herself to wear her gray suit the next day, as well as complex, long-running self-directed behavior change programs in which a person plans and implements one or more contingencies to change his behavior. It is also a functional definition in that the desired change in the target behavior must occur for self-management to be demonstrated.

Self-management is a relative concept. A behavior change program may entail a small degree of self-management or be totally conceived, designed, and implemented by the person. Self-management occurs on a continuum along which the person controls one or all components of a behavior change program. When a behavior change program is implemented by a person or team (e.g., therapist, parent, school staff) on behalf of another (e.g., client, child, student), the external change agent manipulates motivating operations, arranges discriminative stimuli, provides response prompts, delivers differential consequences, and observes and records the occurrence or nonoccurrence of the target behavior. Some degree of self-management is involved whenever a person performs (that is to say, controls) any element of a program that changes his behavior.

It is important to recognize that defining self-management as the personal application of behavior change tactics that results in a desired change in behavior does not explain the phenomenon. Our definition of self-management is descriptive only, and purposively broad. Although self-management tactics can be classified according to their emphasis on a given component of the three- or four-term contingency, or by their structural similarity with a particular principle of behavior (e.g., stimulus control, reinforcement), it is likely that all self-management tactics involve multiple principles of behavior. Thus, when researchers and practitioners describe self-management tactics, they should provide a detailed statement of the procedures used. Behavior analysts should not ascribe the effects of self-management interventions to specific principles of behavior in the absence of an experimental analysis demonstrating such relations. Only through such research will applied behavior analysis develop a more complete understanding of the mechanisms that account for the effectiveness of self-management.[3]

Terminology: Self-Management or Self-Control?

Although *self-management* and *self-control* often appear interchangeably in the behavioral literature, we recommend that *self-management* be used in reference to a person acting "in some way *in order to* change subsequent behavior" (Epstein, 1997, p. 547, emphasis in original). We make this recommendation for three reasons. First, *self-control* is an "inherently misleading" term that implies that the ultimate control of behavior lies within the person (Brigham, 1980). Although Skinner (1953) acknowledged a person could achieve practical control over a given behavior by acting in ways that manipulate variables influencing that behavior, he argued that the controlling behaviors themselves are learned from the person's interactions with the environment.

> A man may spend a great deal of time designing his own life—he may choose the circumstances in which he is to live with great care, and he may manipulate his daily environment on an extensive scale. Such activity appears to exemplify a high order of self-determination. But it is also behavior, and we account for it in terms of other variables in the environment and history of the individual. It is these variables which provide the ultimate control. (p. 240)

In other words, the causal factors for "self-control" (i.e., controlling behaviors) are to be found in a person's experiences with his environment. Beginning with the example of a person setting an alarm clock (controlling behavior) so as to get out of bed at a certain time (controlled behavior) and concluding with an impressive instance of self-control by the main character in Homer's *Odyssey*, Epstein (1997) described the origin of self-control as follows:

> As is true of all operants, any number of phenomena might have produced the [controlling] behavior originally: instructions, modeling, shaping, or generative process (Epstein, 1990, 1996), for example. Its occurrence might have been verbally mediated by a self-generated rule ("I'll bet I'd get up earlier if I set an alarm clock"), and that rule, in turn, might have had any number of origins. Odysseus had his men tie him to his mast (controlling behavior) to lower the probability that he would steer toward the Siren's song (controlled behavior). This was an elegant instance of self-control, but it was entirely instruction driven: Circe had told him to do it. (p. 547)

Second, attributing the cause of a given behavior to self-control can serve as an explanatory fiction. As Baum (2017) pointed out, *self-control* "seems to suggest controlling a [separate] self inside or [that there is] a self inside controlling external behavior. Behavior analysts reject such views as mentalistic. Instead they ask, 'What is the behavior that people call "self-control"?'" (p. 167, words in brackets added). Sometimes *self* is synonymous with *mind* and "not far from the ancient notion of homunculus—an inner person who behaves in precisely the ways necessary to explain the behavior of the outer person in whom he dwells" (Skinner, 1974, p. 121).[4]

Third, laypeople and behavior analysts alike often use *self-control* to refer to a person's ability to delay gratification. In operant terms, this definition of **self-control [as impulse control]** entails responding to achieve a delayed, but larger or higher quality, reinforcer instead of acting to obtain an immediate, less valuable, reinforcer (Rachlin & Greene, 1972; see Box 29.1). Using the same term to refer to a tactic for changing behavior and to a certain type of behavior that may be an outcome of that tactic is confusing and logically faulty. Restricting the use of *self-control* to describe a certain type of behavior (i.e., not acting impulsively) may reduce the confusion caused by using *self-control* as the name for both an independent variable and dependent variable. In this sense, self-control can be a possible goal or outcome of a behavior change program irrespective of whether that behavior is the product of an intervention implemented by an external agent or the subject. Thus, a person can use self-management to achieve, among other things, self-control.

BOX 29.1

Do You Want $10 Now or $100 Next Year?

One way that behavior analysts conceptualize self-control is based on what organisms do when offered choices between rewards of different value and when those rewards can be obtained. Given a choice between a small reward and a larger reward, humans and nonhuman laboratory animals will most often select the larger reward. Offered a choice between an immediate reward and a delayed reward of equal value, organisms tend to choose the immediate reward. Such simple decisions are easy to predict—who doesn't want rewards sooner and have more rather than less?—and require no self-control (Green & Estle, 2003; Odum, 2011).

But things get more difficult to predict, and typical of situations we humans face every day, when the choice is between a smaller, sooner reward (SSR) and a larger, later reward (LLR). In behavioral research, choosing the SSR (e.g., smoking a cigarette) over the LLR (e.g., better health in the future) is considered impulsive behavior. Forgoing the SSR to obtain an LLR reward is viewed as self-control (Ainslie, 1974; Rachlin & Green, 1972).

Delayed rewards, regardless of their significance and magnitude (e.g., enough money for a secure retirement), exert decreasing influence over choice-making behavior as a function of their temporal distance from present circumstances. Both humans and nonhuman laboratory animals discount the value of delayed rewards; the greater the delay to the reward, the greater the discount (i.e., the less value or influence the reward has on current behavior. Behavior analysts call this phenomenon **delay discounting** (or *temporal discounting*) (Odum, 2011; Madden & Johnson, 2010).

Children with attention-deficit hyperactivity disorder (ADHD; Neef et al., 2005), compulsive gamblers (Dixon, Marley, & Jacobs, 2003), heavy smokers (Bickel, Odum, & Madden, 1999), binge drinkers (Vuchinich & Simpson, 1998), people with obesity (Davis, Patte, Curtis, & Reid, 2010), and heroin addicts (Kirby, Petry, & Bickel, 1999) more steeply discount the value of delayed rewards than do comparison groups, suggesting increased likelihood of impulsiveness. Behavior analysts have researched three strategies for decreasing impulsivity: gradually increasing the delay to the LLR, engaging in intervening activities during the delay, and making a commitment response.

Progressive Delay

The progressive delay strategy, first developed to teach "self-control" to pigeons (Ferster, 1953; Mazur & Logue, 1978), essentially shapes preference for LLRs. Subjects initially choose between an SSR and LLR, with each reward available immediately (or after a brief and equal delay, such as a 2-sec delay). After the subject selects the LLR some criterion number of times, the choice trials are repeated with a slight delay to the LLR. The duration of delay to the LLR is gradually increased over sessions.

Schweitzer and Sulzer-Azaroff (1988) reported one of the first applied studies using the progressive delay technique. Subjects were six preschool children described as impulsive by their teachers. Preassessment confirmed the children consistently chose SSRs over LLRs. Following treatment, five of the children chose the LLR more often and three children consistently preferred the LLR at all delay intervals tested. Progressive delay training has reduced impulsive choices by children with ADHD (Neef, Bicard, & Endo, 2001; Neef, Mace, & Shade, 1993), children and adults with autism (Dixon & Cummings, 2001), and adults with developmental disabilities (Dixon & Holcomb, 2000).

Intervening Activities

Engaging in intervening activities during the LLR delay interval may increase self-control. Binder, Dixon, and Ghezzi (2000) required the three children with ADHD to engage in an intervening verbal activity during the delay to the LLR. Two conditions with intervening activities during the delay alternated in multielement fashion. In one condition, the children repeated aloud,

(*continued*)

BOX 29.1 (continued)

"If I wait a little longer, I will get the bigger one." In the other condition, the children named pictures of objects depicted on flash cards. All children exhibited self-control by selecting the delayed rewards during both intervening activity conditions.

A self-control training procedure combining progressive delay and intervening activities has also been evaluated (Dixon et al., 1998; Dixon, Rehfeldt, & Randich, 2003; Dunkel-Jackson, Dixon, & Szekely, 2016). In one study, three adults with autism spectrum disorder were asked, "Do you want a little [reinforcer] now or a lot of [reinforcer] after [doing measuring cups or handwriting]"? (Dunkel-Jackson et al., 2016, p. 707). The delay to the LLR, and thus the duration of task engagement required, was then progressively increased. All three participants chose to engage in the intervening target behavior and earn the LLR on 79%, 77%, and 92% of the intervention trials. This approach not only enhanced self-control—it also increased task engagement.

Commitment Response

Azrin and Powell (1968) conducted a study to limit chain smoking that featured a specially designed cigarette case that locked after the smoker removed a cigarette. By choosing to only smoke cigarettes taken from the case, a participant could not smoke another cigarette for a predetermined amount of time, which increased from 6 minutes to 65 minutes over the course of the study. Rachlin and Green (1972) suggested that a smoker's use of the locking case was a commitment response leading to self-control. Since the smoker is just about to have a cigarette, the value of a future cigarette is negligible, making it easy to shut the case and forgo the next cigarette for a period of time. Similarly, an employee's commitment response of enrolling in a payroll saving plan at the beginning of the month, when money from future paychecks is unavailable (i.e., at a discounted value), is easier than depositing money from a paycheck in hand into a savings account (behavior leading to LLR) instead of impulsively spending it on another dinner out and tickets to a concert (the SSR).

Rachlin (2016) distinguishes between hard commitments to future behavior (snapping shut the locking cigarette case) and soft commitments that create patterns of behavior over time that come into contact with temporally distant contingencies (e.g., I can smoke as many cigarettes as I want today, but I commit to smoking the same number of cigarettes for each of the next 6 days).

To learn more about the experimental analysis, implications, and potential applications of delay discounting, see Critchfield and Kollins (2001); Hursh, Madden, Spiga, DeLeon, and Francisco (2013); and Jacobs, Borrero, and Vollmer (2013).

APPLICATIONS, ADVANTAGES, AND BENEFITS OF SELF-MANAGEMENT

In this section, we identify four basic applications of self-management and describe numerous advantages and benefits that accrue to people who use self-management, practitioners who teach it others, and to society as a whole.

Self-Management Applications

Self-management can help a person be more effective and efficient in daily life, replace bad habits with good ones, accomplish difficult tasks, and achieve personal goals.

Living a More Effective and Efficient Daily Life

Raylene's writing notes to herself and placing her library book on her car seat are examples of self-management techniques for overcoming forgetfulness or lack of organization. Most people use simple self-management techniques, such as making a shopping list before going to the store or creating "to-do" lists, as ways to organize their day; but few probably consider such ordinary efforts to be "self-management." Although many of the most widely used self-management techniques can be considered common sense, people with an understanding of basic principles of behavior can use that knowledge to apply those commonsense techniques more systematically and consistently in their lives.

Breaking Bad Habits and Acquiring Good Ones

Many behaviors we want to do more (or less) often and know that we should do (or not do) are caught in reinforcement traps. Baum (2017) suggested that impulsiveness, bad habits, and procrastination are products of naturally existing *reinforcement traps* in which immediate but smaller consequences have greater influence on our behavior than more significant outcomes that are delayed. Baum gave this description of a reinforcement trap for smoking:

> Acting impulsively leads to a small but relatively immediate reinforcer. The short-term reinforcement for smoking . . . lies in the effects of nicotine and social reinforcers such as appearing grown-up or sophisticated. The trouble with impulsive behavior lies in long-term ill effects. Gradually, over months or years, the bad habit takes its toll in consequences such as cancer, heart disease, and emphysema.
>
> The alternative to impulsiveness, self-control, . . . refraining from smoking, . . . also leads to both short-term and long-term consequences. The short-term consequences are punishing but relatively minor and short-lived: withdrawal symptoms (e.g., headaches) and possibly social discomfort. In the long term, however, . . . refraining from smoking reduces the risk of cancer, heart disease, and emphysema; ultimately, it promotes health. (pp. 168–169)

Reinforcement traps are two-sided contingencies that promote bad habits while simultaneously working against the selection of behavior that is beneficial in the long term. Even though we may know the rules describing such contingencies, the rules are difficult to follow. Malott (1984) described weak rules as those with delayed, incremental, or unpredictable outcomes. An example of a weak rule is *I'd better not smoke or I might get cancer someday and die*. Even though the potential consequence—cancer and death—is a major one, it is far off in the future and not a sure thing even then, which severely limits its effectiveness as a behavioral consequence. So, the rule *don't smoke or you might get cancer* is hard to follow. The deleterious effect of any one cigarette is small, so small that it is not even noticed. Emphysema and lung cancer are likely years and thousands of cigarettes away. How can one more puff hurt?

Self-management provides one strategy for avoiding the deleterious effects of reinforcement traps. A person can use self-management tactics to arrange immediate consequences that will counteract the consequences currently maintaining the self-destructive behavior.

Accomplishing Difficult Tasks

Immediate outcomes that bring a person only infinitesimally closer to a significant, long-term outcome do not control behavior. Many self-management problems are the result of outcomes that are small, but of cumulative significance. Malott (2005a) argued that our behavior is controlled by the outcome of each individual response, not by the cumulative effect of a large number of responses.

> The *natural contingency* between a response and its outcome will be *ineffective* if the outcome of each individual response is too small or too improbable, even though the cumulative impact of those outcomes is significant. So, we have trouble managing our own behavior when each instance of that behavior has only an insignificant outcome, though the outcomes of many repetitions of that behavior will be highly significant. And we have little trouble managing our own behavior, when each instance of that behavior has a significant outcome, though that outcome might be greatly delayed.
>
> Most obese people in the United States know the rule describing that contingency: *If you repeatedly overeat, you will become overweight*. The problem is, knowing the overweight rule does not suppress this one instance of eating that delicious fudge sundae topped with the whipped cream and the maraschino cherry, because a single instance of eating this dessert will cause no significant harm, and will, indeed, taste great. Knowledge of this rule describing the natural contingency exerts little control over gluttony.
>
> But the outcome of our behavior also needs to be probable. Many people have trouble following the buckle-up rule, even though they might get in a serious auto accident as soon as they pull onto the street. Even though the delay between the failure to buckle up and the accident could be only a few seconds, they fail to buckle up because the probability of the accident is so low. However, if they were driving in a dangerous auto race or a dangerous stunt-driving demonstration, they would always buckle up, because the probability of a serious accident would be fairly high. (pp. 516–517)

Just as a person who smokes just one more cigarette or eats another hot fudge sundae cannot detect that lung cancer or obesity is a noticeably closer outcome, graduate student Daryl does not perceive that writing one more sentence brings him closer to his goal of a completed master's thesis. Daryl is capable of writing one more sentence but he often has trouble getting started and working steadily on difficult tasks in which each response produces little or no apparent change in the size of the remaining task. So, he procrastinates.

Applying self-management to this type of performance problem involves designing and implementing one or more contrived contingencies to compete with the ineffective natural contingencies. The self-management contingency provides immediate consequences or short-term outcomes for each response or small series of responses. These contrived consequences increase the occurrence of the target responses, which over time produce the cumulative effects necessary to complete the task.

Achieving Personal Goals

People can use self-management to achieve personal goals, such as learning to play a musical instrument, learning a foreign language, running a marathon, sticking to a schedule of daily yoga sessions (Hammer-Kehoe, 2012), or simply taking some time each day to relax (Harrell, 2012) or listen to enjoyable music (Dams, 2012). For example, a graduate student who wanted to become a better guitarist used self-management to increase the time he practiced scales and chords (Rohn, 2002). He paid a $1 fine to a friend for each day that he failed to practice playing scales and chords for a criterion number of minutes before going to bed. Rohn's self-management program also included several contingencies based on the Premack principle (see Chapter 11). For example, if he practiced scales (ordinarily a low-rate behavior) for 10 minutes, he could play a song (a high-rate behavior).

Advantages and Benefits of Self-Management

A dozen potential advantages and benefits accrue to people who use self-management tactics and to practitioners who teach self-management skills to their clients and students.

Self-Management Can Influence Behaviors Not Accessible to External Change Agents

Self-management can be used to change behaviors that are inaccessible to observation by others. Behaviors such as thoughts of self-doubt, obsessive thoughts, and feelings of depression are private events for which a self-managed treatment approach may be needed (e.g., Kostewicz, Kubina, & Cooper, 2000; Kubina, Haertel, & Cooper, 1994).

Even publicly observable target behaviors are not always emitted in settings and situations accessible to an external change agent. The behavior a person wishes to change may need to be prompted, monitored, evaluated, or reinforced or punished on a day-to-day or even minute-to-minute basis in all of the situations and environments that the person encounters. Most successful smoking, weight loss, exercise, and habit-reversal programs, while planned and directed by others, rely heavily on participants using various self-management techniques while away from the treatment setting. Many behavior change goals brought to clinicians pose the same challenge: How can an effective contingency be arranged that follows the client at all times wherever she goes? For example, target behaviors designed to increase an employee's self-esteem and assertiveness can be identified and practiced in the therapist's office, but implementation of an active contingency in the workplace is apt to require self-management techniques.

External Change Agents Often Miss Important Instances of Behavior

In most education and treatment settings, particularly large-group situations, many important responses go unnoticed by the person responsible for applying behavior change procedures. Classrooms, for example, are especially busy places where desired behaviors by students often go unnoticed because the teacher is engaged with other tasks and students. As a result, students miss opportunities to respond, or they respond and receive no feedback because, in a behavioral sense, the teacher is not there. However, students who have been taught to evaluate their own performance and provide their own feedback in the form of self-delivered rewards and error correction, including seeking assistance and praise from their teacher when needed (e.g., Alber & Heward, 2000; Rouse, Everhart-Sherwood, & Alber-Morgan, 2014), are not dependent on the teacher's direction and feedback for every learning task.

Self-Management Can Promote the Generalization and Maintenance of Behavior Change

A behavior change that (a) continues after treatment has ended, (b) occurs in relevant settings or situations other than the one(s) in which it was learned originally, or (c) spreads to other related behaviors has generality (Baer, Wolf, & Risley, 1968). Important behavior changes without such generalized outcomes must be supported indefinitely by continued treatment.[5] When the student or client is no longer in the setting in which a behavior change was acquired, she may no longer emit the desired response. Particular aspects of the original treatment environment, including the person (teacher, therapist, parent) who administered the behavior change program, may have become discriminative stimuli for the newly learned behavior, enabling the learner to discriminate the presence or absence of certain contingencies across settings. Generalization to nontreatment environments is also hampered when the naturally existing contingencies in those settings do not reinforce the target behavior.

Self-management may overcome challenges to achieving generalized outcomes. Baer and Fowler (1984) posed and answered a pragmatic question related to the problem of promoting the generalization and maintenance of newly learned skills.

> What behavior change agent can go with the student to every necessary lesson, at all times, to prompt and reinforce every desirable form of the behavior called for by the curriculum? The student's own "self" can always meet these specifications. (p. 148)

A Small Repertoire of Self-Management Skills Can Control Many Behaviors

A person who learns a few self-management tactics can control a potentially wide range of behaviors. For example, self-monitoring—observing and recording one's own behavior—has been used to increase on-task behavior (Clemons, Mason, Garrison-Kane, & Wills, 2016), academic productivity and accuracy (e.g., Rock, 2005), homework completion and accuracy (Falkenberg & Barbetta, 2013), productivity at work (e.g., Christian & Poling, 1997), physical activity (Hayes & Van Camp, 2015), and reciprocal social conversations (Koegel, Park, & Koegel, 2014), and to decrease target behaviors such as cigarette smoking (Foxx & Brown, 1979) and stereotypy or tics (Crutchfield, Mason, Chambers, Wills, & Mason, 2015; Fritz, Iwata, Rolider, Camp, & Neidert, 2012; Tiger, Fisher, & Bouxsein, 2009).

People with Diverse Abilities Can Learn Self-Management Skills

People of wide-ranging ages and cognitive abilities have used self-management tactics successfully. Preschoolers (e.g., Sainato, Strain, Lefebvre, & Rapp, 1990), typically developing students from primary grades through high school (Busacca, Moore, & Anderson, 2015), students with learning disabilities (McDougall et al., 2017), students with emotional and behavior disorders (Bruhn, McDonald, & Kreigh, 2015), children with autism (Carr, Moore, & Anderson, 2014; Southall & Gast, 2011), and children and adults with intellectual and developmental disabilities (Grossi & Heward, 1998; Reynolds, Gast, & Luscre, 2014) have all used self-management successfully. Even college professors are capable of improving their performance with self-management (Malott, 2005a)!

Some People Perform Better Under Self-Selected Tasks and Performance Criteria

Participants in some studies have performed better with self-selected work tasks and consequences than under contingencies determined by others (e.g., Baer, Tishelman, Degler, Osnes, & Stokes, 1992; Olympia, Sheridan, Jenson, & Andrews, 1994; Parsons, Reid, Reynolds, & Bumgarner, 1990). For example, in three brief experiments with a single student, Lovitt and Curtiss (1969) found pupil-selected rewards and contingencies to be more effective than teacher-selected performance standards. In the first phase of Experiment 1, the subject, a 12-year-old boy in a special education classroom, earned a teacher-specified number of minutes of free time based on the number of math and reading tasks he completed correctly. In the next phase of the study, the student was allowed to specify

the number of correct math and reading items needed to earn each minute of free time. During the final phase of the study, the original teacher-specified ratios of academic production to reinforcement were again in effect. The student's median academic response rate during the self-selected contingency phase was 2.5 correct responses per minute (math and reading tasks reported together), compared to correct rates of 1.65 and 1.9 in the two teacher-selected phases.

An experiment by Dickerson and Creedon (1981) used a yoked-control (Sidman, 1960) and between-group comparison experimental design with 30 second- and third-graders and found that pupil-selected standards resulted in significantly higher academic production of both reading and math tasks. Both the Lovitt and Curtiss (1969) and the Dickerson and Creedon studies demonstrated that self-selected reinforcement contingencies *can* be more effective than teacher-selected contingencies. However, research in this area also shows that simply letting children determine their own performance standards does not guarantee high levels of performance; some studies have found that children select too-lenient standards when given the opportunity (Felixbrod & O'Leary, 1973, 1974). But interestingly, in a study of student-managed homework teams by Olympia and colleagues (1994), although students assigned to teams who could select their own performance goals chose more lenient accuracy criteria than the 90% accuracy criterion required of students on the teams that worked toward teacher-selected goals, the overall performance of the teams with student-selected goals was slightly higher than that of the teams working toward teacher-selected goals. More research is needed to determine the conditions under which students will self-select and maintain appropriate standards of performance.

People with Good Self-Management Skills Contribute to Efficient and Effective Group Environments

The overall effectiveness and efficiency of any group of people who share a working environment is limited when a single person is responsible for monitoring, supervising, and providing feedback for everyone's performance. Use of self-management skills by students, teammates, band members, or employees without having to rely on teachers, coaches, band directors, or managers for every task, improves the overall performance of the entire group. In the classroom, for example, teachers traditionally assume full responsibility for setting performance criteria and goals for students, evaluating students' work, delivering consequences for those performances, and managing students' social behavior. The time required for these functions is considerable. Students who can self-score their work, provide their own feedback using answer keys or self-checking materials (Bennett & Cavanaugh, 1998; Goddard & Heron, 1998), and behave appropriately without teacher supervision enable their teacher to attend to other aspects of the curriculum and perform other instructional duties (Mitchem, Young, West, & Benyo, 2001; Saddler & Good, 2006).

When Hall, Delquadri, and Harris (1977) conducted observations of elementary classrooms and found low levels of active student response, they surmised that higher rates of academic productivity might actually be punishing to teachers. Even though considerable evidence links high rates of student responding with academic achievement (e.g., Hattie, 2009; MacSuga-Gage & Simonsen, 2015), generating more student responses in most classrooms results in the teacher grading more papers. The time savings produced by students with even the simplest self-management skills can be significant. In one study, a class of five elementary special education students completed as many arithmetic problems as they could during a 20-min session each day (Hundert & Batstone, 1978). When the teacher graded the papers, an average of 50.5 minutes was required to prepare and conduct the session and to score and record each student's performance. When the students scored their own papers, the total teacher time needed to conduct the arithmetic period averaged 33.4 minutes.

Teaching Students Self-Management Skills Provides Meaningful Practice in Other Curriculum Areas

When students learn to define and measure behavior and to graph, evaluate, and analyze their own responses, they practice various math and science skills in a relevant way. When students are taught how to conduct self-experiments such as using A-B designs to evaluate their self-management projects, they receive meaningful practice in logical thinking and scientific methodology (Marshall & Heward, 1979; Moxley, 1998).

Self-Management Is an Ultimate Goal of Education

When asked to identify what education should accomplish for its students, most people—educators and laypeople alike—include in their reply the development of independent, self-directed citizens capable of behaving appropriately and constructively without the supervision of others. John Dewey (1939), one of America's most influential educational philosophers, said, "the ideal aim of education is the creation of self-control" (p. 75). A student's ability to be self-directed, with the teacher as a guide or facilitator, and his ability to evaluate his own performance have long been considered cornerstones of humanistic education.

As Lovitt (1973) observed, the fact that systematic instruction of self-management skills is *not* a regular part of most schools' curricula is a paradox because "one of the expressed objectives of the educational system is to create individuals who are self-reliant and independent" (p. 139). Although self-control is expected and valued by society, it is seldom addressed directly by the school curriculum. Teaching self-management as an integral part of the curriculum requires that students learn a fairly sophisticated sequence of skills (e.g., Marshall & Heward, 1979; McConnell, 1999). However, systematic teaching of self-management skills, if successful, is well worth the effort required; students will have an effective means of dealing with situations in which there is little, if any, external control.

Self-Management Benefits Society

Self-management serves two important functions for society (Epstein, 1997). First, citizens with self-management skills are more likely to fulfill their potential and make greater contributions to society. Second, self-management helps people behave

in ways (e.g., purchasing fuel-efficient automobiles, taking public transportation) that forgo immediate reinforcers that are correlated with devastating outcomes so long deferred that only future generations will experience them (e.g., depletion of natural resources and climate change). People who have learned self-management techniques that help them save resources, recycle, and burn less fossil fuel help create a better world for everyone. Self-management skills can provide people with the means of meeting the well-intentioned but difficult-to-follow rule, *Think globally but act locally*.

Self-Management Helps a Person Feel Free

Baum (2017) noted that people caught in reinforcement traps who recognize the contingency between their addictive behavior, impulsiveness, or procrastination and the eventual consequence that is likely to occur as a result of that behavior are unhappy and do not feel free. However, "Someone who escapes from a reinforcement trap, like someone who escapes from coercion, feels free and happy. Ask anyone who has kicked an addiction" (p. 193).

Ironically, a person who skillfully uses self-management techniques derived from a scientific analysis of behavior–environment relations predicated on an assumption of determinism is more likely to feel free than is a person who believes that her behavior is a product of free will. In discussing the apparent disconnect between philosophical determinism and self-control, Epstein (1997) compared two people: one who has no self-management skills and one who is very skilled at managing his own affairs.

> First, consider the individual who has no self-control. In Skinner's view, such a person falls prey to all immediate stimuli, even those that are linked to delayed punishment. Seeing a chocolate cake, she eats it. Handed a cigarette, she smokes it. . . . She may make plans, but she has no ability to carry them out, because she is entirely at the mercy of proximal events. She is like a sailboat blowing uncontrollably in a gale . . .
>
> At the other extreme, we have a skillful self-manager. He, too, sets goals, but he has ample ability to meet them. He has skills to set dangerous reinforcers aside. He identifies conditions that affect his behavior and alters them to suit him. He takes temporally remote possibilities into account in setting his priorities. The wind is blowing, but he sets the boat's destination and directs it there.
>
> These two individuals are profoundly different. The first is being controlled in almost a linear fashion by her immediate environment. The second is, in a nontrivial sense, controlling his own life. . . .
>
> The woman who lacks self-control skills *feels controlled*. She may believe in free will (in fact, in our culture, it's a safe bet that she does) but her own life is out of control. A belief in free will only exacerbates her frustration. She should be able to *will* herself out of any jam, but "willpower" proves to be highly unreliable. In contrast, the self-manager feels that *he is in control*. Ironically, like Skinner, he may believe in determinism, but he not only feels that he is in control, he is in fact exercising considerably more control over his life than our impulsive subject. (p. 560, emphasis in original)

Self-Management Feels Good

A final, but by no means trivial, reason for learning self-management is that being in control of one's life feels good. A person who arranges her environment in purposeful ways that support and maintain desirable behaviors not only will be more productive than she would otherwise be but also will feel good about herself. Seymour (2002) noted her personal feelings about a self-management intervention she implemented with the objective of running for 30 minutes 3 times per week.

> The guilt I had been feeling for the $2\frac{1}{2}$ past years had been burning strong within me. [As a former scholarship softball player] I used to exercise 3 hours per day (6 days per week) because that exercise paid for my schooling. . . . But now the contingencies were missing, and my exercising had gone completely down the tubes. Since the start of my project, I've run 15 out of the 21 times; so I've gone from 0% during baseline to 71% during my intervention. I'm happy with that. Because the intervention's successful, the guilt I have felt for the past years has been lifted. I have more energy and my body feels new and strong. It's amazing what a few contingencies can do to improve the quality of life in a big way. (pp. 7–12)

ANTECEDENT-BASED SELF-MANAGEMENT TACTICS

In this and the three sections that follow, we describe some of the many self-management tactics that behavioral researchers and clinicians have developed. Although no standard set of labels or classification scheme for self-management tactics has emerged, the techniques are often presented according to their relative emphasis on antecedents or consequences for the target behavior. An antecedent-based self-management tactic features the manipulation of events or stimuli antecedent to the target (controlled) behavior. Sometimes grouped under general terms such as *environmental planning* (Bellack & Hersen, 1977; Thoresen & Mahoney, 1974) or *situational inducement* (Martin & Pear, 2015), antecedent-based approaches to self-management encompass a wide range of tactics, including the following:

- Manipulating motivating operations to make a desired (or undesired) behavior more (or less) likely.
- Providing response prompts.
- Performing the initial steps of a behavior chain to ensure being confronted later with a discriminative stimulus that will evoke the desired behavior.
- Removing the materials required for an undesired behavior.
- Limiting an undesired behavior to restricted stimulus conditions.
- Dedicating a specific environment for a desired behavior.

Manipulate Motivating Operations

A motivating operation (MO) is an environmental condition or event that (a) alters the effectiveness of some stimulus, object, or event as a reinforcer and (b) alters the current rate of all behaviors that have been reinforced by that stimulus, object, or event in the past (see Chapter 16). An MO that increases the effectiveness of a reinforcer and has an evocative effect on behaviors that have produced that reinforcer is called an *establishing operation* (EO); an MO that decreases the effectiveness of a reinforcer and has an abative (i.e., weakening) effect on behaviors that have produced that reinforcer is called an *abolishing operation* (AO).[6]

The general strategy for incorporating an MO into a self-management intervention is to behave in a way (controlling behavior) that creates a certain state of motivation that, in turn, increases (or decreases, as desired) subsequent occurrences of the target behavior (controlled behavior). Imagine you have been invited to dinner at the house of someone known to be a first-rate cook. You want to obtain maximum enjoyment from what promises to be a special meal but are worried you will not be able to eat everything. Skipping lunch (controlling behavior) creates an EO that would increase the likelihood of your being able to enjoy the evening meal from appetizer through dessert (controlled behavior). Conversely, eating a meal just before going grocery shopping (controlling behavior) could serve as an AO that decreases the momentary value of ready-to-eat foods with high sugar and fat content as reinforcers and results in purchasing fewer such items at the supermarket (controlled behavior).

Provide Response Prompts

Creating stimuli that later function as extra cues and reminders for desired behavior is a simple and often effective self-management technique. Response prompts can take a wide variety of forms (e.g., visual, auditory, textual, symbolic), be permanent for regularly occurring events (e.g., entering "Put trash out tonight!" each Thursday on one's calendar), or be one-time affairs, as when Raylene wrote, "gray suit today" on a sticky-note and attached it to her closet door where she was sure to see it when dressing for work in the morning. Dieters often put pictures of overweight people, or perhaps even unattractive pictures of themselves, on the refrigerator door, near the ice cream, in the cupboard—anyplace where they might look for food. Seeing the pictures may evoke other controlling responses, such as moving away from the food, phoning a friend, going for a walk, or marking a point on an "I Didn't Eat It!" chart.

The same object can be used as a generic response prompt for various behaviors, such as putting a rubber band around the wrist as a reminder to do a certain task at some later time. This form of response prompt, however, will be ineffective if the person is unable later to remember what specific task the physical cue was intended to prompt. Engaging in some self-instruction when "putting on" a generic response prompt may help the person recall the task to be completed when seeing the cue later. For example, one of the authors of this text uses a small carabiner he calls a "memor-ring" as a generic response prompt. When the thought of an important task that can only be completed in another setting later in the day occurs to him, he clips the carabiner to a belt loop or the handle of his briefcase and "activates the memor-ring" by stating the target task to himself 3 times (e.g., "borrow Nancy's journal at Arps Hall," "borrow Nancy's journal at Arps Hall," "borrow Nancy's journal at Arps Hall"). When he sees the "memor-ring" in the relevant setting later, it usually functions as an effective prompt for completing the targeted task.

Self-arranged cues can also be used to prompt the occurrence of a behavior the person wants to repeat in a variety of settings and situations. In this case, the person might seed his environment with supplemental response prompts. For example, a father who wants to increase the number of times he interacts with and praises his children might place colorful sticky-notes printed with the word ZAP! (see Figure 29.2) in various

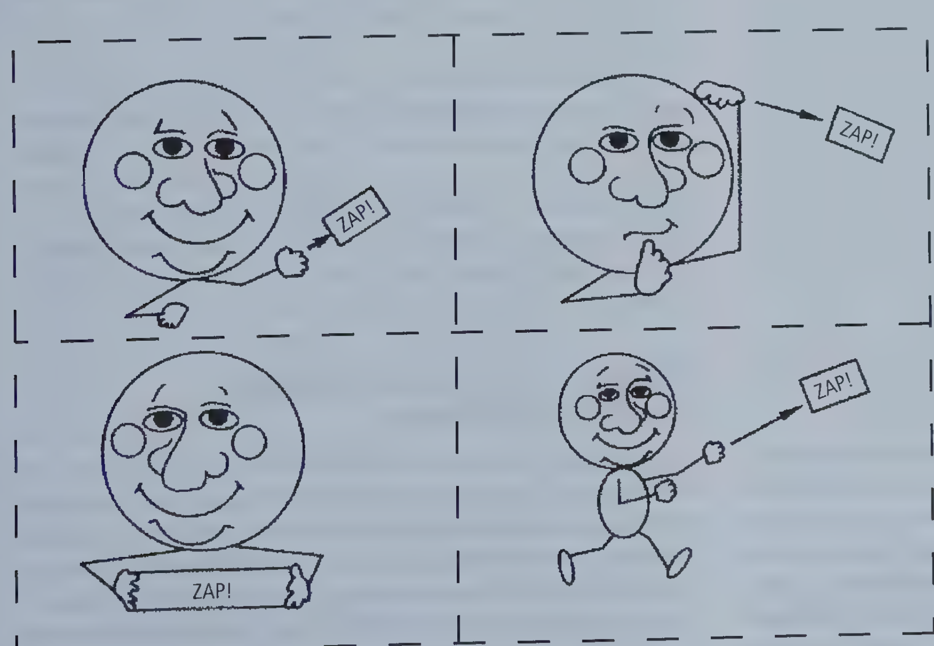


Figure 29.2 Contrived cues used by parents to prompt self-selected parenting behaviors.

Originally developed by Charles Novak. From *Working with Parents of Handicapped Children*, by W. L. Heward, J. C. Dardig, and A. Rossett, 1979, p. 26, Columbus, OH: Charles E. Merrill.

places around the house where he will see them routinely—on the microwave, on the TV remote control, as a bookmark. Each time he sees a ZAP! cue, the father is reminded to make a comment or ask a question of his child, or to look for some positive behavior for which he can provide attention and praise.

Smartphones, tablets, watches, and other portable digital devices can provide self-delivered audio and visual response prompts wherever target behavior should be emitted (Cullen & Alber-Morgan, 2015; Mechling, 2011). Individuals with autism and other disabilities have used digital devices to prompt acquisition and continued performance across a wide range of daily-living (Cannella-Malone, Brooks, & Tullis, 2013; Cannella-Malone, Chan, & Jimenez, 2017) and vocational tasks (Cullen, Alber-Morgan, Simmons-Reed, & Izzo, 2017).

Perform the Initial Steps of a Behavior Chain

Another self-management tactic featuring manipulation of antecedent events involves behaving in a manner that ensures your future behavior contacts a discriminative stimulus (S^D) that reliably evokes the target behavior. Although operant behavior is selected and maintained by its consequences, the moment-to-moment occurrence of many behaviors is influenced by the presence or absence of discriminative stimuli.

Many tasks consist of response chains. Each response in a chain produces a change in the environment that serves both as a conditioned reinforcer for the response that preceded it and as an S^D for the next response in the chain (see Chapter 23). Successful completion of a behavior chain requires each component response to be evoked by its related S^D. By performing part of a behavioral chain (the self-management response) at one point in time, a person changes her environment so that she will be confronted later with an S^D that will evoke the next response in the chain and will lead to the completion of the task (the self-managed response). Skinner (1983b) provided an excellent example of this tactic.

> Ten minutes before you leave your house for the day you hear a weather report: It will probably rain before you return. It occurs to you to take an umbrella (the sentence means quite literally what it says: The behavior of taking an umbrella occurs to you), but you are not yet able to execute it. Ten minutes later you leave without the umbrella. You can solve that kind of problem by executing as much of the behavior as possible when it occurs to you. Hang the umbrella on the doorknob, or put it through the handle of your briefcase, or in some other way start the process of taking it with you. (p. 240)

Remove Items Necessary for an Undesired Behavior

Altering the environment to make an undesired behavior less likely or, better yet, impossible to emit is another antecedent-based self-management tactic. The smoker who discards her cigarettes and the dieter who removes all cookies and chips from his house, car, and office, have, for the moment at least, effectively controlled their smoking and junk food eating. Although other self-management efforts will probably be needed to refrain from seeking and obtaining the harmful material, ridding oneself of the items needed to engage in an undesired behavior is a good start.

Limit Undesired Behavior to Restricted Stimulus Conditions

A person may be able to decrease an undesired behavior by limiting the setting or stimulus conditions where he engages in the behavior. To the extent that the restricted situation acquires stimulus control over the target behavior and access to the situation is infrequent or otherwise not reinforcing, the target behavior will occur less frequently. Imagine a man who habitually touches and scratches his face (he is aware of this bad habit because his wife often nags him about it and asks him to stop). In an effort to decrease his face touching, the man decides that he will do two things. First, whenever he becomes aware that he is engaging in face touching, he will stop; second, at any time he can go to a bathroom and touch and rub his face for as long as he wants.

Nolan (1968) reported the case of a woman who used restricted stimulus conditions in an effort to quit smoking. The woman noted she most often smoked when other people were around or when she watched television, read, or lay down to relax. It was decided she would smoke only in a prescribed place, and she would eliminate from that place the potentially reinforcing effects of other activities. She designated a specific chair as her "smoking chair" and positioned the chair so she could not watch TV or easily engage in conversation while sitting in it. She asked family members not to approach her or talk to her while she was in her smoking chair. She followed the procedure faithfully, and her smoking decreased from a baseline average of 30 cigarettes per day to 12 cigarettes per day. Nine days after beginning the program, the woman decided to try to reduce her smoking even further by making her smoking chair less accessible. She put the chair in the basement, and her smoking decreased to five cigarettes per day. A month after beginning the smoking chair program, the woman had quit smoking completely.

Goldiamond (1965) used a similar tactic to help a man who continually sulked when interacting with his wife. The husband was instructed to limit his sulking behavior to his "sulking stool," which was placed in the garage. The man went to the garage whenever he felt like sulking, sat on his stool and sulked for as long as he wanted, and left when he was finished sulking. The man found that his sulking decreased considerably when he had to sit on a stool in the garage to do it.

Dedicate a Specific Environment for a Desired Behavior

A person may achieve some degree of stimulus control over a behavior that requires diligence and concentration by reserving or creating an environment where he will only engage in that behavior. Students have improved their study habits and professors their scholarly productivity when they have selected a specific place in which to study, free of other distractions, and have not engaged in any other behaviors, such as daydreaming

or letter writing, in that place (Goldiamond, 1965). Skinner (1981b) suggested this kind of stimulus control strategy when he offered the following advice to aspiring writers:

> Equally important are the conditions in which the behavior occurs. A convenient place is important. It should have all the facilities needed for the execution of writing. Pens, typewriters, recorders, files, books, a comfortable desk and chair. . . . Since the place is to take control of a particular kind of behavior, you should do nothing else there at any time. (p. 2)

People who do not have the luxury of devoting a certain environment to a single activity can contrive a special stimulus that can be turned on and off in the multipurpose setting, as in this example provided by Watson and Tharp (2014):

> A man had in his room only one table, which he had to use for a variety of activities, such as writing letters, paying bills, watching TV, and eating. But when he wanted to do concentrated studying or writing, he always pulled his table away from the wall and sat on the other side of it. In that way, sitting on the other side of his table became the cue associated only with concentrated intellectual work. (p. 168)

Most students have a personal computer they use for their academic work but also for writing and reading personal e-mails, conducting family business, browsing the Internet, playing games, shopping online, and so on. In a manner similar to the man who sat on the other side of his table when studying and writing, a student might display a particular background on the computer's desktop that signals that only academic work is to be done on the computer. For example, when he sits down to study or write, a student could replace the standard display on his computer desktop, a photo of his dog, let's say, with a solid green background. The green desktop signals that only scholarly work will be done. Over time the "scholarly work" desktop may acquire a degree of stimulus control over the desired behavior. If the student wants to quit working and use the computer for something else, he must first change the background on the desktop. Doing so before he has fulfilled his time or productivity goal for a work session may generate some guilt that can be escaped by returning to work.

This tactic can also be used to increase a desired behavior that is not being emitted at an acceptable rate because of competing undesired behaviors. In one case, an adult with insomnia reportedly went to bed about midnight but did not fall asleep until 3:00 or 4:00 AM. Instead of falling asleep, the man would worry about mundane problems and turn on the television. Treatment consisted of instructing the man to go to bed when he felt tired, but he was not to stay in bed if he was unable to sleep. If he wanted to think about his problems or watch television, he could do so, but he was to get out of bed and go into another room. When he once more felt sleepy, he was to return to bed and try again to go to sleep. If he still could not sleep, he was to leave the bedroom again. The man reported getting up 4 or 5 times a night for the first few days of the program, but within 2 weeks he went to bed, stayed there, and fell asleep.

SELF-MONITORING

Self-monitoring has been the subject of more research and clinical applications than any other self-management strategy. **Self-monitoring** (also called *self-recording* or *self-observation*) is a procedure whereby a person systematically observes his behavior and records the occurrence or nonoccurrence of a target behavior. Originally conceived as a clinical assessment technique for collecting data on behaviors that only clients could observe and record (e.g., eating, smoking, fingernail biting), self-monitoring became a major therapeutic intervention in its own right because of the reactive effects it often produced. As discussed in Chapters 3 and 4, *reactivity* refers to the effects an assessment or measurement procedure has on a person's behavior. In general, the more obtrusive the observation and measurement method, the greater the likelihood of reactivity. Maximum obtrusiveness exists when the person observing and recording the target behavior is the subject of the behavior change program, and reactivity is very likely.

Although reactivity caused by a researcher's measurement system represents uncontrolled variability in a study and must be minimized as much as possible, the reactive effects of self-monitoring are welcome from a clinical perspective. Not only does self-monitoring often change behavior, but also the change is typically in the educationally or therapeutically desired direction.

Behavior analysts have taught adults to use self-monitoring to decrease smoking (Lipinski, Black, Nelson, & Ciminero, 1975; McFall, 1977), increase physical activity (Kurti & Dallery, 2013) and calorie expenditure (Donaldson & Normand, 2009), and stop biting their nails (Maletzky, 1974). Students with and without disabilities have used self-monitoring to be on task more often in the classroom (Legge, DeBar, & Alber-Morgan, 2010; Wills & Mason, 2014), decrease disruptive behaviors (Kamps, Conklin, & Wills, 2015; Martella, Leonard, Marchand-Martella, & Agran, 1993), improve reading comprehension (Crabtree, Alber-Morgan, & Konrad, 2010), increase performance in a variety of academic subject areas (Holifield, Goodman, Hazelkorn, & Heflin, 2010; Wolfe, Heron, & Goddard, 2000), and complete homework assignments (Falkenberg & Barbetta, 2013). Classroom teachers have used self-monitoring to increase their use of specific praise statements during classroom instruction (Hager, 2012; Simonsen, MacSuga, Fallon, & Sugai, 2013) and the fidelity with which they conduct instructional procedures (Belfiore, Fritts, & Herman, 2008; Pinkelman & Horner 2016; Plavnick, Ferreri, & Maupin, 2010).

In one of the first published accounts of self-monitoring by children, Broden, Hall, and Mitts (1971) analyzed the effects of self-recording on the behavior of two eighth-grade students. Liza was earning a D– in history and exhibited poor study behavior during the lecture format class. Using 10-sec momentary time sampling for 30 minutes each school day, an observer seated in the back of the classroom (Liza was not told she was being observed) produced a 7-day baseline showing that Liza exhibited study behaviors (e.g., faced the teacher, took notes when appropriate) an average of 30% of the observation intervals, despite two conferences with the school counselor in which she promised to "really try." Prior to the eighth session,

the counselor gave Liza a piece of paper with three rows of 10 squares and directed her to record her study behavior "when she thought of it" during her history class sessions. Some aspects of study behavior were discussed at this time, including a definition of what constituted studying.

> Liza was instructed to take the slip to class each day and to record a "+" in the square if she was studying or had been doing so for the last few minutes, and a "−" if she was not studying at the time she thought to record. Sometime before the end of the school day she was to turn it in to the counselor. (p. 193)

Figure 29.3 shows the results of Liza's self-monitoring. Her level of study behavior increased to 78% (as measured by the independent observer) and stayed at approximately that level during Self-Recording 1, quickly decreased to baseline levels when self-recording was terminated during Baseline 2, and averaged 80% during Self-Recording 2. During Self-Recording Plus Praise, Liza's teacher attended to her whenever he could during history class and praised her for study behavior whenever possible. Liza's level of study behavior during this condition increased to 88%.

The bottom graph in Figure 29.3 shows the number of times per session the observer recorded teacher attention toward Liza. That teacher attention did not correlate in any apparent way with changes in Liza's study behavior during the first four phases of the experiment suggests that teacher attention—which often exerts a powerful influence on student behavior—was not a confounding variable and that the improvements in Liza's study behavior most likely can be attributed to the self-recording procedure. However, the effects of self-monitoring may have been confounded by the fact that Liza turned her self-recording slips in to the counselor before the end of each school day, and at weekly student–counselor conferences she received praise from the counselor for recording slips with a high percentage of plus marks.

In the second experiment reported by Broden and colleagues (1971), self-monitoring was used by Stu, an eighth-grade student who talked out in class "continually." An observer recorded Stu's talk-outs during both parts of a math class that met before and after lunch. Self-recording was begun first during the before-lunch portion of the math class. The teacher handed Stu a slip of paper on which was marked a 2-inch by 5-inch rectangular box and the instruction, "Put a mark every time you talk out" (p. 196). Stu was told to use the recording slip and turn it in after lunch. At the top of the recording slip was a place for the student's name and the date. No other instructions were given to Stu, nor were any consequences or other contingencies

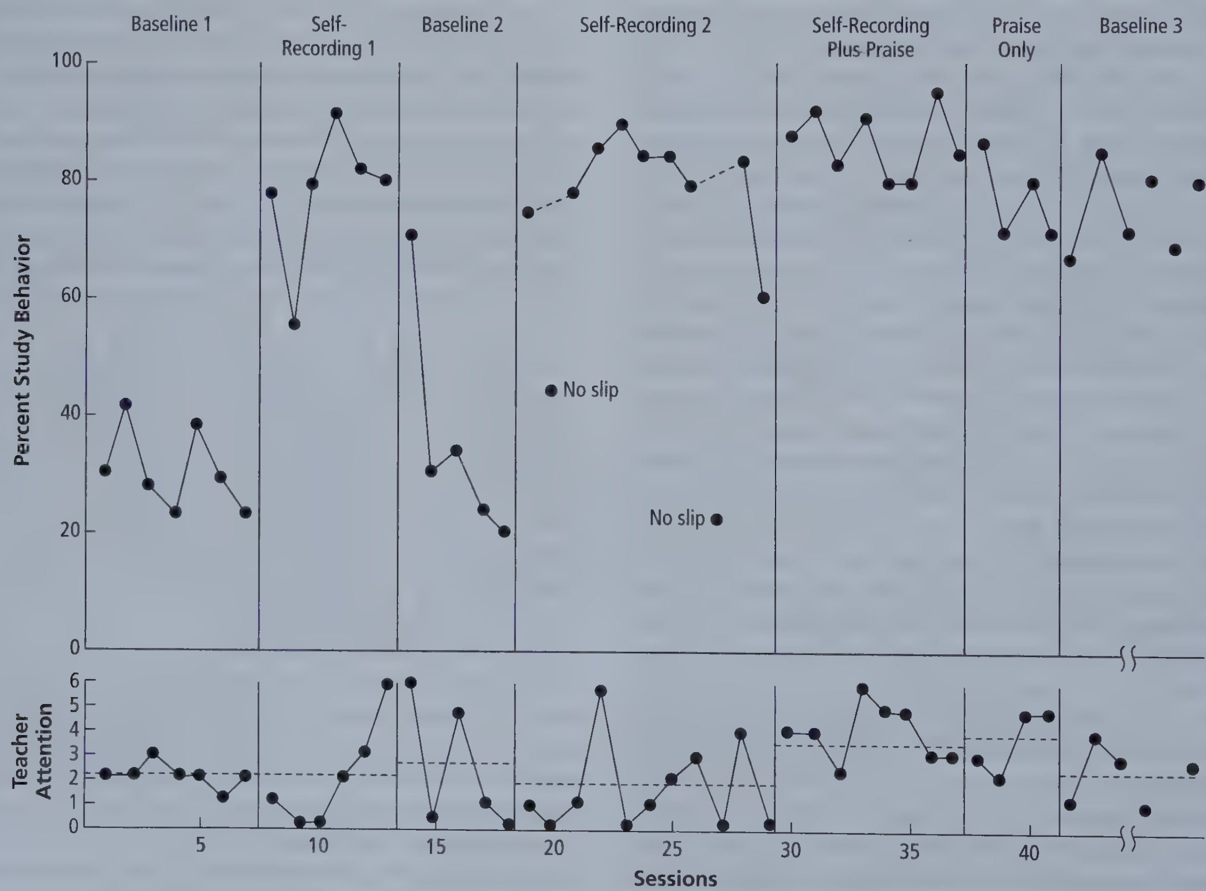


Figure 29.3 Percentage of observed intervals in which an eighth-grade girl paid attention in class.

Republished with permission of Society for the Experimental Analysis of Behavior, Inc., from The Effect of Self-Recording on the Classroom Behavior of Two Eighth-Grade Students, M. Broden, R. V. Hall, and B. Mitts, 1971, *Journal of Applied Behavior Analysis*, *4*, p. 194. Permission conveyed through Copyright Clearance Center, Inc.

applied to his behavior. Self-recording was later implemented during the second half of the math class. During the pre-lunch portion of the math class, Stu's baseline rate of 1.1 talk-outs per minute was followed by a rate of 0.3 talk-outs per minute when he self-recorded. After lunch Stu talked out at an average rate of 1.6 times per minute before self-recording and 0.5 times per minute during self-recording.

The combination reversal and multiple baseline design showed a clear functional relation between the self-monitoring procedure and Stu's reduced talk-outs. During the final phase of the experiment, however, Stu talked out at rates equal to initial baseline levels, even though self-recording was in effect. Broden and colleagues (1971) suggested that the lack of effect produced by self-recording during the final phase may have resulted from the fact that "no contingencies were ever applied to differential rates of talking out and the slips thus lost their effectiveness" (p. 198). It is also possible that the initial decreases in talk-outs attributed to self-recording were confounded by Stu's expectations of some form of teacher reinforcement.

As can be seen from these two experiments, it is extremely difficult to isolate self-monitoring as a straightforward, "clean" procedure; self-monitoring almost always entails other contingencies. Nevertheless, the various and combined procedures that constitute self-monitoring are effective in changing behavior.

Self-Evaluation

Self-monitoring is often combined with goal setting and self-evaluation. A person using **self-evaluation** (also called *self-assessment*) compares her performance with a predetermined goal or standard. Self-evaluation can be aided by self-graphing self-collected data (Kasper-Ferguson & Moxley, 2002; Stolz, Itoi, Konrad, & Alber-Morgan, 2008). For example, Grossi and Heward (1998) taught four restaurant trainees with developmental disabilities to select productivity goals, self-monitor their work, and self-evaluate their performance against a competitive productivity standard (i.e., the typical rate at which nondisabled restaurant employees perform the task in competitive settings). Job tasks in this study included scrubbing pots and pans, loading a dishwasher, busing and setting tables, and mopping and sweeping floors.

Self-evaluation training with each trainee took approximately 35 minutes across three sessions and consisted of five parts: (1) rationale for working faster (e.g., finding and keeping a job in a competitive setting), (2) goal setting (trainee was shown a simple line graph of his baseline performance compared to the competitive standard and prompted to set a productivity goal), (3) use of timer or stopwatch, (4) self-monitoring and how to chart self-recorded data on a graph that showed the competitive standard with a shaded area, and (5) self-evaluation (comparing his own work against the competitive standard and making self-evaluative statements, e.g., "I'm not in the shaded area. I need to work faster." or "Good, I'm in the area."). When a trainee met his goal for 3 consecutive days, a new goal was selected. At that time, all four trainees had selected goals within the competitive range.

The work productivity of all four trainees increased as a function of the self-evaluation intervention. Figure 29.4 shows the results for one of the participants, Chad, a 20-year-old man with mild cognitive disabilities, cerebral palsy, and a seizure disorder. For the 3 months prior to the study, Chad had been training at two work tasks at the dishwashing job station: scrubbing cooking pots and pans and racking dishes (a six-step behavior chain for loading an empty dishwasher rack with dirty dishes). Throughout the study, pot scrubbing was measured during 10-min observation sessions both before and after the daily lunch operation. Racking dishes was measured using a stopwatch to time to the nearest second how long a trainee took to load a rack of dishes; 4 to 8 racks per session were timed during the busiest hour of the lunch shift. After each observation session during baseline, Chad received feedback on the accuracy and quality of his work. No feedback was given during baseline on the productivity of his performance.

During baseline, Chad scrubbed an average of 4.5 pots and pans every 10 minutes, and none of his 15 baseline trials was within the competitive range of 10 to 15 pots. Chad's pot scrubbing rate increased to a mean of 11.7 pots during self-evaluation, with 76% of 89 self-evaluation trials at or above the competitive range. During baseline, it took Chad a mean of 3 minutes 2 seconds to load one rack of dishes, and 19% of his 97 baseline trials were within the competitive range of 1 to 2 minutes. During self-evaluation, Chad's time to load one rack improved to a mean of 1 minute 55 seconds, and 70% of the 114 self-evaluation trials were within the competitive range. At the end of the study, three of the four trainees stated that they liked to time and record their work performance. Although one trainee stated it was "too stressful" to time and record his work, he said that self-monitoring helped him show other people that he could do the work.

Participants who have self-evaluated their behavior as captured on video have improved their posture while sitting at computer work stations (Sigurdsson & Austin, 2008) and their performance of complex yoga postures (Downs, Miltenberger, Biedronski, & Witherspoon, 2015).

Self-Monitoring with Reinforcement

Self-monitoring is often part of an intervention package that includes feedback and reinforcement for meeting either self- or teacher-selected goals (e.g., Bulla & Frieder, 2017; Christian & Poling, 1997; Olympia et al., 1994). The reinforcer may be self-administered or teacher delivered. For example, Koegel, Koegel, Hurley, and Frea (1992) taught four children with autism between the ages of 6 and 11 to obtain their own reinforcers after they had self-recorded a criterion number of appropriate responses to questions from others (e.g., "Who drove you to school today?").

Teacher-delivered reinforcement was a component of a self-monitoring intervention used by Martella and colleagues (1993) to help Brad, a 12-year-old student with mild intellectual disabilities, reduce the number of negative statements he made during classroom activities (e.g., "I hate this #*%@! calculator," "Math is crappy"). Brad self-recorded negative statements during two class periods, charted the number on a graph, and then compared his count with the count obtained by a student teacher. If Brad's self-recorded data agreed with the teacher at

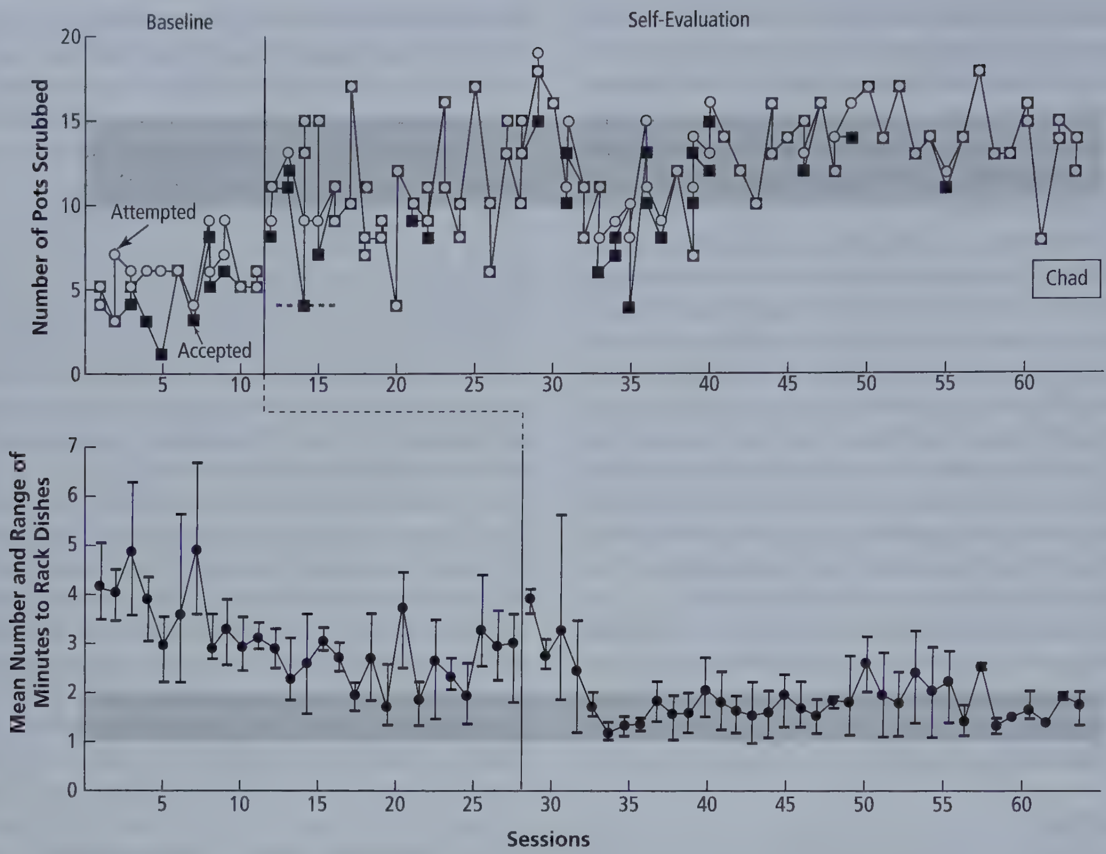


Figure 29.4 Number of pots scrubbed in 10 minutes and the mean and range of minutes to load a rack of dishes in the dishwasher during baseline and self-evaluation conditions. Shaded bands represent the competitive performance range of restaurant workers without disabilities. The dashed horizontal line represents Chad's self-selected goal; no dashed line means that Chad's goal was within the competitive range. Vertical lines through data points on the bottom graph show the range of Chad's performance across 4 to 8 trials.

a level of 80% or higher, he received his choice from a menu of "small" reinforcers (items costing 25 cents or less). If Brad's self-recorded data agreed with the teacher's data and were at or below a gradually declining criterion level for four consecutive sessions, he was allowed to choose a "large" (more than 25 cents) reinforcer (see Figure 30.16).

Why Does Self-Monitoring Work?

The behavioral mechanisms that account for the effectiveness of self-monitoring are not fully understood. Some behavior theorists suggest that self-monitoring is effective in changing behavior because it evokes self-evaluative statements that serve either to reinforce desired behaviors or to punish undesired behaviors. Cautela (1971) hypothesized that a child who records on a chart that he has completed his chores may emit covert verbal responses (e.g., "I am a good boy") that serve to reinforce chore completion. Malott (1981) suggested that self-monitoring improves performance because of what he called *guilt control.* Self-monitoring less-than-desirable behavior produces covert guilt statements that can be avoided by improving one's performance. That is, the target behavior is strengthened through negative reinforcement by escape and avoidance of the guilty feelings that occur when one's behavior is "bad."

Descriptions of the self-monitoring techniques used by two famous authors apparently agreed with Malott's guilt control hypothesis. Novelist Anthony Trollope, writing in his 1883 autobiography, stated:

> When I have commenced a new book, I have always prepared a diary, divided into weeks, and carried on for the period which I have allowed myself for the completion of the work. In this I have entered, day by day, the number of pages I have written, so that if at any time I have slipped into idleness for a day or two, the record of that idleness has been there staring me in the face, and demanding of me increased labour, so that the deficiency might be supplied. . . . I have allotted myself so many pages a week. The average number has been about 40. It has been placed as low as 20, and has risen to 112. And as a page is an ambiguous term, my page has been made to contain 250 words; and as words, if not watched, will have a tendency to straggle, I have had every word counted as I went. . . . There has ever been the record before me, and a week passed with an

insufficient number of pages has been a blister to my eye and a month so disgraced would have been a sorrow to my heart. (from Wallace, 1977, p. 518)

Guilt control also helped to motivate the legendary Ernest Hemingway. Novelist Irving Wallace (1977) reported the following extract from an article by George Plimpton about the self-monitoring technique used by Hemingway.

> He keeps track of his daily progress—"so as not to kid myself"—on a large chart made out of the side of a cardboard packing case and set up against the wall under the nose of a mounted gazelle head. The numbers on the chart showing the daily output of words differ from 450, 575, 462, 1250, back to 512, the higher figures on days Hemingway put in extra work so he won't feel guilty spending the following day fishing on the Gulf Stream. (p. 518)

Exactly which principles of behavior are operating when self-monitoring changes the target behavior is not known because much of the self-monitoring procedure consists of private, covert behaviors. In addition to the problem of access to these private events, self-monitoring is usually confounded by other variables. Self-monitoring is often part of a self-management package in which contingencies of reinforcement, punishment, or both, are included, either explicitly (e.g., "If I run 10 miles this week, I can go to the movies.") or implicitly (e.g., "I've got to show my record of calories consumed to my wife."). Regardless of the principles of behavior involved, however, self-monitoring is often an effective procedure for changing one's behavior.

Self-Monitoring Guidelines and Procedures

Practitioners should consider the following suggestions when implementing self-monitoring with their clients.

Provide Materials That Make Self-Monitoring Easy

If self-monitoring is difficult, cumbersome, or time-consuming, at best it will be ineffective and disliked by the participant, and at worst it may have negative effects on the target behavior. Provide participants with materials and devices that make self-monitoring as simple and efficient as possible. All of the low-tech devices and procedures for measuring behavior described in Chapter 4 (e.g., paper and pencil, wrist counters, hand-tally counters, timers, stopwatches) can be used for self-monitoring. For example, a teacher wanting to praise students at least 10 times during a class period could put 10 pennies in her pocket prior to the start of class. Each time she praises a student's behavior, she moves one coin to another pocket.

Recording forms for most self-monitoring situations should be very simple. Self-recording forms consisting of little more than a series of boxes or squares are often effective (e.g., Broden et al., 1971). At various intervals, the participant might write a plus or minus, circle yes or no, or mark an X on a smiling face or sad face, in a kind of momentary time sampling procedure; or he could record a count of the number of responses made during a just-completed interval.

Recording forms can be created for self-monitoring specialized tasks or a chain of behaviors. Dunlap and Dunlap (1989) taught students with learning disabilities to self-monitor the steps they used to solve subtraction problems with regrouping. Each student self-monitored his work by recording a plus or a minus next to each step on an individualized checklist of steps (e.g., "I underlined all the top numbers that were smaller than the bottom."; "I crossed out only the number next to the underlined number and made it one less." [p. 311]) designed to prompt the student from committing specific types of errors.

Lo (2003) taught elementary students who were at risk for behavioral disorders to use the form shown in Figure 29.5 to self-monitor whether they worked quietly, evaluated their work, and followed a prescribed sequence for obtaining teacher

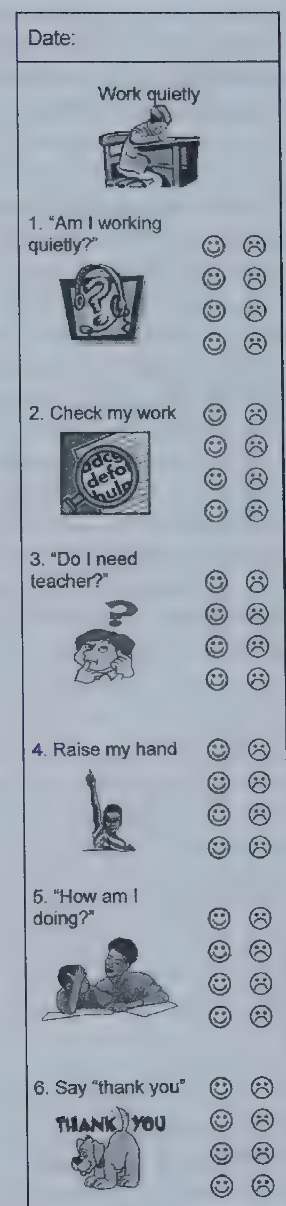


Figure 29.5 Form used by elementary school students to self-monitor whether they worked quietly and followed a prescribed sequence for recruiting teacher assistance.

From "*Functional Assessment and Individualized Intervention Plans: Increasing the Behavioral Adjustment of Urban Learners in General and Special Education Settings*" by Y. Lo, 2003. Unpublished doctoral dissertation. Columbus, OH: The Ohio State University. Reprinted by Permission.

assistance during independent seatwork activities. The form, which was taped to each student's desk, functioned as a reminder to students of the expected behaviors and as a device on which they self-recorded those behaviors.

"Countoons" are self-monitoring forms that illustrate the contingency with a series of cartoon-like frames. A countoon reminds young children of not only what behavior to record but also what consequences will follow if they meet specified performance criteria. Daly and Ranalli (2003) created six-frame countoons that enable students to self-record an inappropriate behavior and an incompatible appropriate behavior. In the countoon shown in Figure 29.6, Frames F1 and F4 show the student doing her math work, appropriate behavior that is counted in Frame F5. The criterion number of math problems to meet the contingency—in this case, 10—is also indicated in Frame F5. Frame F2 shows the student talking with a friend, the inappropriate behavior to be counted in F3. The student must not chat more than 6 times to meet the contingency. The "What Happens" frame (F6) depicts the reward the student will earn by meeting both parts of the contingency. Daly and Ranalli have provided detailed steps for creating countoons and using them to teach self-management skills to children.

Self-monitoring forms can also be designed to enable self-recording of multiple tasks across days. For example, teachers working with secondary students might find the Classroom Performance Record (CPR) developed by Young and colleagues (1991) an effective way to assist students with monitoring their assignments, completion of homework, points earned, and citizenship scores. The form also provides students with information about their current status in the class, their likely semester grade, and tips for improving their performance.

Self-monitoring interventions to improve fitness can be aided by smartwatches and other digital devices that automatically count steps, measure heart rate, and record variables such as time and distance (e.g., Hayes & Van Camp, 2015; Valbuena, Miltenberger, & Solley, 2015). In one study, college student athletes self-monitored their class attendance and arrival time by text messaging their academic counselors (Bicard, Lott, Mills, Bicard, & Baylot-Casey, 2012).

Provide Supplementary Cues or Prompts

Although some self-recording devices—a recording form taped to a student's desk, a notepad for counting calories carried in a dieter's shirt pocket, a golf counter on a teacher's wrist—serve as continual reminders to self-monitor, additional prompts or cues to self-monitor are often helpful. Researchers and practitioners have used a variety of auditory, visual, and tactile stimuli as cues and prompts for self-monitoring.

Auditory prompts in the form of prerecorded signals or tones have been widely used to cue self-monitoring in classrooms (e.g., Todd, Horner, & Sugai, 1999). For example, in a study by Glynn, Thomas, and Shee (1973), second-graders placed a checkmark in a series of squares if they thought they were on task at the moment they heard a tape-recorded beep. A total of 10 beeps occurred at random intervals during a 30-min class period. Using a similar procedure with a digital device, Gulchak (2008) taught an 8-year-old boy diagnosed with emotional and behavior disorders to use a stylus to check *Yes* or *No* under the heading "Are You On-Task?" on a handheld computer when the computer's alarm chimed at 10-min intervals.

Ludwig (2004) used visual cues on the classroom board to prompt kindergarten children to self-record their productivity during a morning seatwork activity in which they wrote answers on individual worksheets to items, problems, or questions that the teacher had printed on a large whiteboard on the classroom wall. The work was divided into 14 sections by topic and covered a variety of curriculum areas (e.g., spelling, reading comprehension, addition and subtraction problems, telling time, money).

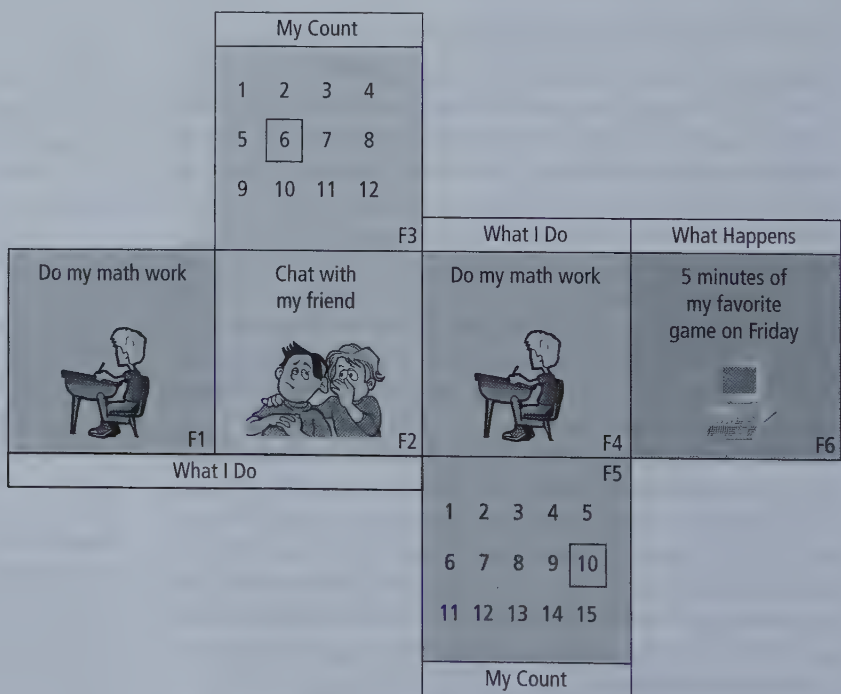


Figure 29.6 Example of a countoon that can be taped to a student's desk as a reminder of target behaviors, the need to self-record, and the consequence for meeting the contingency.

Republished with permission of Council for Exceptional Children, from Using Countoons to Teach Self-Monitoring Skills, P. M. Daly and P. Ranalli, 2003, *Teaching Exceptional Children*, *35*(5), p. 32. Permission conveyed through Copyright Clearance Center, Inc.

At the end of each section, the experimenter drew a smiley face with a number from 1 to 14. The numbered smiley faces drawn on the board corresponded to 14 numbered smiley faces on each student's self-monitoring card.

Digital devices can provide unobtrusive visual prompts to self-monitor. Clemons and colleagues (2016) programmed a handheld tablet's touchscreen to flash "Am I on-task?" followed by the choices "yes" and "no" at 30-sec or 1-min intervals. Three high school students with varying disabilities exhibited improved levels of on-task behavior when using the system in general and self-contained special education classrooms.

Tactile prompts can also be used to signal self-recording opportunities. For example, the *MotivAider(R)* (www.habitchange.com) is a small electronic device that vibrates at fixed or variable time intervals programmed by the user. This device is excellent for signaling people to self-monitor or perform other self-management tasks (e.g., Legge et al., 2010) or to prompt practitioners to attend to the behavior of students or clients (Flaute, Peterson, Van Norman, Riffle, & Eakins, 2005).

Prompts to self-monitor, whatever their form, should be as unobtrusive as possible so they do not disrupt the participant or others. As a general rule, practitioners should provide frequent prompts to self-monitor at the beginning of a self-management intervention and gradually reduce their number as the participant becomes more proficient and skilled with self-monitoring.

Self-Monitor the Most Important Dimension of the Target Behavior

Self-monitoring is measurement. As we saw in Chapter 4, there are various dimensions by which a given behavior can be measured. What dimension of the target behavior should be self-monitored? A person should self-monitor the dimension that, should desired changes in its value occur, would yield the most direct and significant progress toward his self-management goal. For example, a man wanting to lose weight by eating less food could measure the number of bites he takes throughout the day (count), the number of bites per minute at each meal (rate), how much time elapses from the time he sits down at the table to when he takes the first bite (latency), how long he pauses between bites (interresponse time), and/or how long his meals last (duration). Although daily measures of each of these dimensions would provide some quantitative information about the man's eating, none of the dimensions would be as directly related to his goal as would the total number of calories he consumes each day.

Numerous studies have examined the question of whether students should self-monitor their on-task behavior or their academic performance or productivity. Some researchers have found self-monitoring of attention more effective (Harris, Friedlander, Saddler, Frizzelle, & Graham, 2005; Kneedler & Hallahan, 1981), others have reported advantages for self-monitoring performance (Lam, Cole, Shapiro, & Bambara, 1994; Maag, Reid, & DiGangi, 1993; Reid, 1996), and some have found the two procedures equally effective (Carr & Punzo, 1993; Shimabukuro, Pratter, Jenkins, & Edelen-Smith, 2000). Although self-monitoring either attention or performance results in increased on-task behavior, students tend to complete more academic responses when self-monitoring productivity than when self-recording on-task behavior (e.g., Maag et al., 1993; Lam et al., 1994; Reid & Harris, 1993). Additionally, most students prefer self-monitoring academic performance to self-recording on-task behavior (e.g., "You could say the words while you counted them." [Harris et al., 2005, p. 151]).

In general, we recommend teaching students to self-monitor a measure of academic productivity (e.g., number of problems or items attempted, number correct) rather than whether they are on task. This is because increasing on-task behavior, whether by self-monitoring or by contingent reinforcement, does not necessarily result in a collateral increase in productivity (e.g., Marholin & Steinman, 1977; McLaughlin & Malaby, 1972; Wolfe et al., 2000). By contrast, when productivity is increased, improvements in on-task behavior almost always occur. However, a student whose persistent off-task and disruptive behaviors create problems for him or others in the classroom may benefit more from self-monitoring on-task behavior, at least initially.

Self-Monitor Early and Often

In general, each occurrence of the target behavior should be self-recorded as soon as possible. However, the act of self-monitoring a behavior the person wants to increase should not disrupt "the flow of the behavior" (Critchfield, 1999). Self-monitoring behaviors that produce natural or contrived response products (e.g., answers on academic worksheets, words written) can be conducted after the session by permanent product measurement (see Chapter 4).

Relevant aspects of some target behaviors can be self-monitored even before the behavior has occurred. Self-recording a response early in a behavior chain that leads to an undesired behavior the person wants to decrease may be more effective in changing the target behavior in the desired direction than recording on the terminal behavior in the chain. For example, Rozensky (1974) reported that when a woman who had been a heavy smoker for 25 years recorded the time and place she smoked each cigarette, there was little change in her rate of smoking. She then began recording the same information each time she noticed herself beginning the chain of responses that led to smoking: reaching for her cigarettes, removing a cigarette from the pack, and so on. She stopped smoking within a few weeks of self-monitoring in this fashion.

Generally, a person should self-monitor more often at the beginning of a behavior change program. If and as performance improves, self-monitoring can occur less often. For example, Rhode, Morgan, and Young (1983) had students with behavioral disorders begin self-evaluating (on a 0- to 5-point scale) the extent to which they followed classroom rules and completed academic work at 15-min intervals throughout the school day. Over the course of the study, the self-evaluation intervals were increased gradually, first to every 20 minutes, then every 30 minutes, then once per hour. Eventually, the self-evaluation cards were withdrawn and students orally reported their self-evaluations. In the final self-evaluation condition, the students verbally self-evaluated their compliance with classroom rules and academic work on an average of every 2 days (i.e., a variable ratio [VR] 2-day schedule).

Reinforce Accurate Self-Monitoring

Some researchers have reported little correlation between the accuracy of self-monitoring and its effectiveness in changing the behavior being recorded (e.g., Kneedler & Hallahan, 1981; Marshall, Lloyd, & Hallahan, 1993). It appears that accurate self-monitoring is neither a sufficient nor a necessary condition for behavior change. For example, Hundert and Bucher (1978) found that, even though students became highly accurate in self-scoring their arithmetic, their performance on the arithmetic itself did not improve. In contrast, in the Broden and colleagues (1971) study, both Liza and Stu's behavior improved even though their self-recorded data seldom matched the data from independent observers. Nevertheless, accurate self-monitoring is desirable, especially when participants are using their self-recorded data as the basis for self-evaluation or self-administered consequences.

Although some studies have shown that young children can accurately self-record their behavior in the absence of external contingencies for accuracy (e.g., Ballard & Glynn, 1975; Glynn et al., 1973), other researchers have reported low agreement between the self-reported data and the data collected by independent observers (Kaufman & O'Leary, 1972; Turkewitz, O'Leary & Ironsmith, 1975). One factor that seems to affect the accuracy of self-scoring is the use of self-reported scores as a basis for reinforcement. Santogrossi, O'Leary, Romanczyk, and Kaufman (1973) found that, when children self-evaluated their own work and those self-produced evaluations determined levels of token reinforcement, their self-monitoring accuracy deteriorated over time. Similarly, Hundert and Bucher (1978) found that students who previously had accurately self-scored their arithmetic assignments greatly exaggerated their scores when higher scores resulted in points that could be exchanged for prizes.

Rewarding children for producing self-recorded data that match the data of an independent observer and spot-checking students' self-scoring reports can increase the accuracy of self-monitoring by young children. Drabman, Spitalnik, and O'Leary (1973) used these procedures in teaching children with behavior disorders to self-evaluate their own classroom behavior.

> Now something different is going to happen. If you get a rating within plus or minus one point of my rating, you can keep all your points. You lose all your points if you are off my rating by more than plus or minus one point. In addition, if you match my rating exactly you get a bonus point. (O'Leary, 1977, p. 204)

After the children demonstrated that they could evaluate their own behavior reliably, the teacher began checking only 50% of the children's self-ratings by pulling names from a hat at the end of the period, then 33% of the children, then 25%, then 12%. During the last 12 days of the study, she did not check any child's self-rating. Throughout this period of reduced checks and finally no checks at all, the children continued to self-evaluate accurately. Rhode and colleagues (1983) used a similar faded matching technique.

Additional examples and techniques for teaching self-monitoring to children can be found in Joseph and Konrad (2009); McCollow, Curiel, Davis, and Sainato (2016); and Rafferty (2010).

SELF-ADMINISTERED CONSEQUENCES

Arranging for specific consequences to follow occurrences (or nonoccurrences) of one's behavior is a fundamental strategy of self-management. In this section, we review a variety of tactics people have used to self-reinforce and self-punish. First, we briefly examine some of the conceptual issues raised by the concept of "self-reinforcement."

Is Self-Reinforcement Possible?

Skinner (1953) pointed out that *self-reinforcement* should not be considered synonymous with the principle of operant reinforcement.

> The place of operant reinforcement in self-control is not clear. In one sense, all reinforcements are self-administered since a response may be regarded as "producing" its reinforcement, but "reinforcing one's own behavior" is more than this. . . . Self-reinforcement of operant behavior presupposes that the individual has it in his power to obtain reinforcement but does not do so until a particular response has been emitted. This might be the case if a man denied himself all social contacts until he had finished a particular job. Something of this sort unquestionably happens, but is it operant reinforcement? It is roughly parallel to the procedure in conditioning the behavior of another person. But it must be remembered that the individual may at any moment drop the work in hand and obtain the reinforcement. We have to account for his not doing so. It may be that such indulgent behavior has been punished—say, with disapproval—except when a piece of work has just been completed. (pp. 237–238)

In his discussion of self-reinforcement, Goldiamond (1976a) continued with Skinner's example and stated that the fact that a person "does not cheat, but engages in the task, cannot be explained simply by resort to *his* self-reinforcement by social contact, which he makes available contingent on his finishing the job" (p. 510). In other words, the variables influencing the controlling response—in this case abstaining from social contact until the job is complete—must still be accounted for; simply citing self-reinforcement as the cause is an explanatory fiction.

The issue is not whether procedures labeled as "self-reinforcement" often work to change behavior in ways that resemble the defining effect of reinforcement; they do. However, a careful examination of instances of self-reinforcement reveals that something more than, or different from, a straightforward application of positive reinforcement is involved (e.g., Brigham, 1980; Catania, 2013; Goldiamond, 1976a, 1976b). As a technical term, *self-reinforcement* (as also, *self-punishment*) is a misnomer, and the issue is not merely a semantic one, as some writers have suggested (e.g., Mahoney, 1976). Assigning the effectiveness of a behavior change tactic to a well-understood principle of behavior when something other than, or in addition to, that principle is operating overlooks relevant variables that may hold the key to a fuller understanding of the tactic. Considering the analysis of a behavioral episode complete once it has been identified as a case of self-reinforcement precludes further search for other relevant variables.

We agree with Malott (2005a, 2012; Malott & Shane, 2014), who argued that performance-management contingencies, whether designed and implemented by oneself or by others, are best viewed as rule-governed analogs of reinforcement and punishment contingencies because the response-to-consequence delay is too great.[7] Malott (2005a) provided the following examples of self-management contingencies as analogs to negative reinforcement and punishment contingencies:

> [Consider the] contingency, *Every day I eat more than 1,250 calories, I will give someone else $5 to spend frivolously*, like a despised roommate or despised charity, though for many of us, the loss of $5 to a beloved charity is sufficiently aversive that the thought of it will punish overeating. [This] contingency is an analog to a penalty [punishment] contingency, an analog because the actual loss of the $5 will normally occur more than 1 minute after exceeding the 1,250 calorie limit. Such analog penalty contingencies are effective in decreasing undesirable behavior. And to increase desirable behavior, an analog to an avoidance [negative reinforcement] contingency works well: *Every day I exercise for one hour, I avoid paying a $5 fine*. But, if you have not finished your hour by midnight, you must pay the $5. (p. 519, emphasis in original, words in brackets added)

Self-Administered Consequences to Increase Behavior

A person can increase the target behavior in a self-management program by applying contingencies that are analogs to positive and negative reinforcement. Self-selected consequences can be delivered directly (the self-manager accesses the consequence herself) or indirectly (another person provides the consequence).

Self-Management Analogs of Positive Reinforcement

A variety of target behaviors have improved when students determined the number of tokens, points, or minutes of free time earned based on self-assessment of their performance (Glynn, 1970; Koegel et al., 1992; Olympia et al., 1994). In a self-experiment, Craig (2010) decreased the number of times he bit his fingernails by awarding himself points using a token-based differential reinforcement of other behavior (DRO) procedure. A wristwatch with a countdown timer signaled 1-hr DRO intervals. The effects of interventions featuring self-administered rewards are difficult to evaluate because they are typically confounded by other variables. In a review of self-reinforcement research, Jones, Nelson, and Kazdin (1977) identified participants' previous history, criterion setting, self-monitoring, self-evaluation, external contingencies on the self-reinforcing response, external contingencies on the target behavior, and actual or suspected surveillance as likely confounding variables.

A study by Ballard and Glynn (1975) controlled for self-monitoring as a contributor to the effects of self-reinforcement. Third-graders were taught after a baseline condition to self-score and self-record several aspects of their writing—number of sentences, number of describing words, and number of action words. Self-monitoring had no effect on any of the variables measured, even though the students handed in their counting sheets with their writing each day. The children were then given a notebook in which to record their points, which could be exchanged at the rate of one point per minute for each student's choice of activities during an earned-time period each day. The self-reinforcement procedure resulted in large increases in each of the three dependent variables.

In studies on self-recruited reinforcement, students are taught to periodically self-evaluate their work and then show it to their teacher and request feedback or assistance (Alber, Heward, & Hippler, 1999; Craft, Alber, & Heward, 1998; Rouse et al., 2014). In a sense, students administer their own reinforcer by recruiting the teacher's attention, which often results in praise and other forms of reinforcement (see Alber & Heward, 2000 for a review).

Todd and colleagues (1999) taught an elementary student to use a self-management system that included self-monitoring, self-evaluation, and self-recruited reinforcement. Kyle was a 9-year-old boy diagnosed with a learning disability and receiving special education services in reading, math, and language arts. Kyle's individualized education program (IEP) also included several objectives for problem behavior (e.g., disrupting independent and group activities, teasing and taunting classmates, and saying sexually inappropriate comments). The classroom teacher's goal setting and daily evaluation had proven ineffective. An "action team" conducted a functional assessment (see Chapter 27) and designed a support plan that included the self-management system.

Kyle was taught how to use the self-management system during two 15-min one-on-one training sessions in which he practiced self-recording numerous role-played examples and nonexamples of on-task and off-task behavior and learned appropriate ways to recruit teacher attention and praise. For self-monitoring, Kyle listened with a single earplug to a 50-min cassette tape on which 13 checkpoints (e.g., "check one," "check two") had been prerecorded on a variable interval (VI) 4-min schedule (ranging from 3- to 5-min intervals between checkpoints). Each time Kyle heard a checkpoint, he marked a plus (if he had been working quietly and keeping his hands, feet, and objects to himself) or a zero (if he had been teasing peers and/or not working quietly) on a self-recording card.

Todd and colleagues (1999) described the recruitment of teacher praise and how Kyle's special program was integrated into the reinforcement system for the entire class:

> Each time Kyle marked three pluses on his card he raised his hand (during instruction) or walked up to the teacher (during group project time) and requested feedback on his performance. The teacher acknowledged Kyle's good work and placed a mark on his self-monitoring card to show where he should begin a new count of three pluses. In addition to these within-session contingencies, Kyle could earn a self-manager sticker at the end of each class period if he had no more than two zeroes for the period. Stickers were earned by all students in the class for appropriate behavior and were pooled weekly for class rewards. Because these stickers were combined, Kyle's stickers were valued by all the students and provided opportunities for positive peer attention. (p. 70)

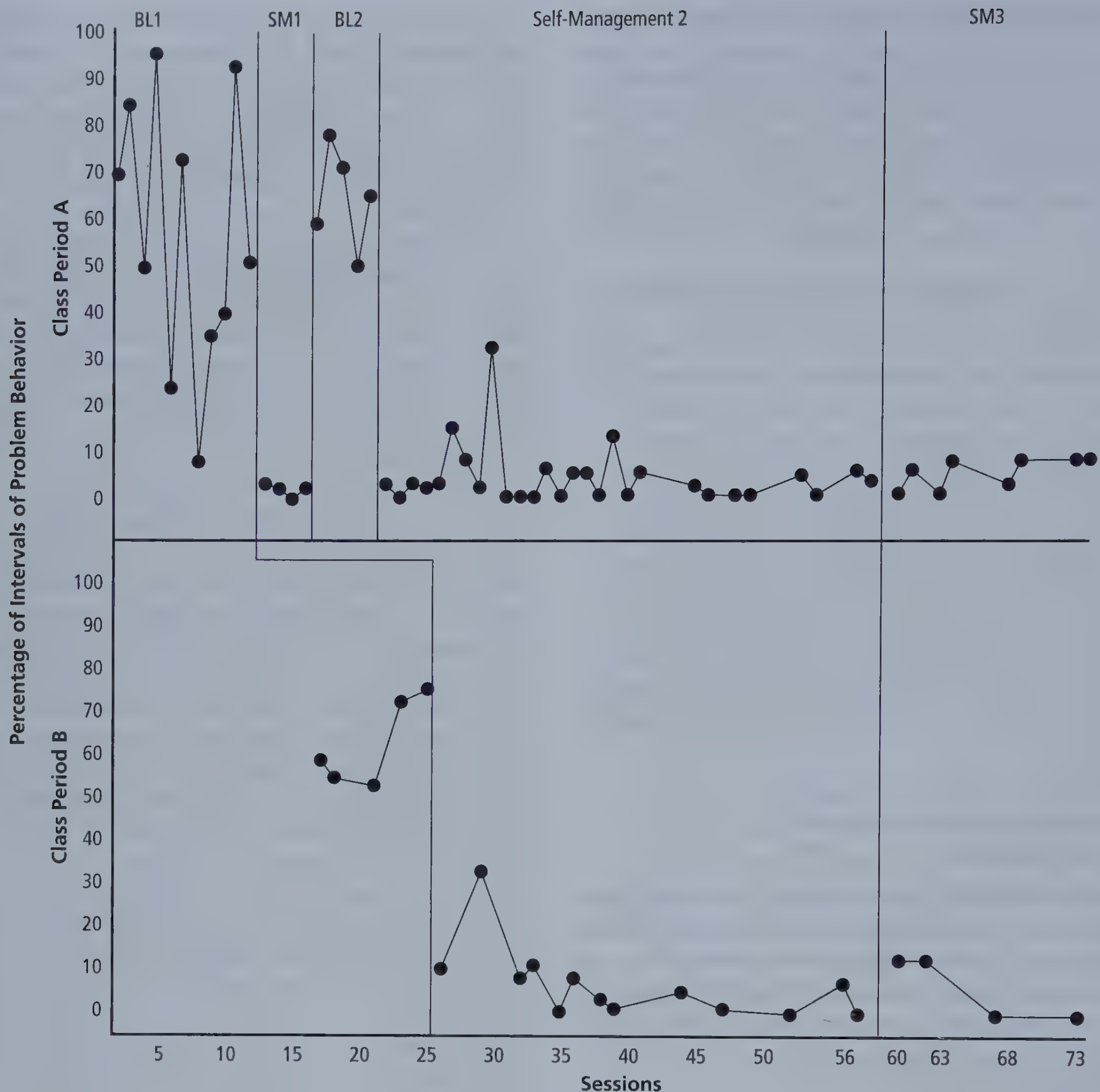


Figure 29.7 Problem behaviors by a 9-year-old boy during 10-min probes across two class periods during baseline (BL) and self-management (SM) conditions.

Republished with permission of Pro-Ed, Inc., from Self-Monitoring and Self-Recruited Praise: Effects on Problem Behavior, Academic Engagement, and Work Completion in a Typical Classroom, A. W. Todd, R. H. Horner, and G. Sugai, 1999, *Journal of Positive Behavior Interventions*, *1*, p. 71. Permission conveyed through Copyright Clearance Center, Inc.

Procedures for the second self-management phase (SM2) were the same as for the first phase; during the third phase (SM3), Kyle used a 95-min tape on which 16 checkpoints were prerecorded on a VI 5-min schedule (ranging from 4 to 6 minutes). When Kyle used the self-management system, the percentage of intervals in which he engaged in problem behavior was much lower than baseline (BL) levels (see Figure 29.7). Large increases in on-task behavior and completion of academic work were also noted. An important outcome of this study was that Kyle's teacher praised his behavior more often. Significantly, the self-management intervention was initiated in Class Period B (bottom graph in Figure 29.7) at the teacher's request because she had noted an immediate and "dramatic change" in Kyle's performance when he began self-monitoring and recruiting teacher praise in Class Period A. This outcome is strong evidence for the social validity of the self-management intervention.

Self-Management Analogs of Negative Reinforcement

Many successful self-management interventions involve self-determined escape and avoidance contingencies analogous to negative reinforcement. Most of the case studies featured in Malott's (2012) book on self-management, *I'll Stop Procrastinating When I Get Around to It*, feature escape and avoidance contingencies in which emitting the target behavior enabled the person to avoid an aversive event. The table below provides examples:

Target Behavior/Objective	Self-Management Contingency
Make a daily journal entry of interesting things that happened so as to have events to include in a weekly letter to her parents.	Each day Kelly failed to make a journal entry, she had to do a friend's chores, including doing the dishes and laundry. (Garner, 2012)
Four $\frac{1}{2}$-hr exercise sessions (e.g., mountain bike, aerobics, strength building) per week; 3 sessions must be during the week and 1 session on the weekend.	Pin $5 bill to the corkboard on the refrigerator and give it to my roommate at the end of the week if the exercise goal remains unmet. (Haroff, 2012)
Practice guitar for half an hour each day of the week before 11:00 PM.	Do 50 sit-ups at 11:00 PM on Sundays for each day that I did not practice during the week. (Knittel, 2012)

At stickK.com individuals can design personalized "commitment contracts" in which they specify behavior change goals, set dates for completion, select referees to verify self-reports of task completion, and escrow a self-determined amount of money to lose if the terms of the contract are not met. In response to the uneasiness some people may feel about using negative reinforcement to control their behavior, Malott (2012) made the following case for building "pleasant aversive control" into self-management programs.

> Aversive control doesn't have to be aversive! . . . Here's what I think you need in order to have a pleasant aversive control procedure. You need to make sure that the aversive consequence, the penalty, is small. And you need to make sure the penalty is usually avoided—the avoidance response needs to be one that the person will readily make, most of the time, as long as the avoidance procedure is in effect.
>
> Our everyday life is full of such avoidance procedures, and yet they don't make us miserable. You don't have an anxiety attack every time you walk through a doorway, even though you might hurt yourself if you don't avoid bumping into the doorjamb. And you don't break out in a cold sweat every time you put the leftovers in the refrigerator, and thereby avoid leaving them out to spoil. So that leads us to the Self-Management Rule: Don't hesitate to use an avoidance procedure to get yourself to do something you want to do anyway. Just be sure the aversive outcome is as small as possible (but not so small it's ineffective) and the response is easy to make. (p. 80)

Self-Administered Consequences to Decrease Behavior

Self-administered consequences analogous to positive punishment or negative punishment can decrease the occurrence of undesirable behavior.

Self-Management Analogs of Positive Punishment

A person can decrease the rate of an undesired behavior by following each occurrence with the onset of painful stimulation or aversive activity. Mahoney (1971) reported a case study in which a man beleaguered by obsessive thoughts wore a heavy rubber band around his wrist. Each time he experienced an obsessive thought, he snapped the rubber band, delivering a brief, painful sensation to his wrist. A 15-year-old girl who had compulsively pulled her hair for $2\frac{1}{2}$ years, to the point of creating bald spots, also used contingent self-delivered snaps of a rubber band on her wrist to stop her habit (Mastellone, 1974). Another woman stopped her hair pulling by performing 15 sit-ups each time she pulled her hair or had the urge to do so (MacNeil & Thomas, 1976).

Self-administering a positive practice overcorrection procedure also qualifies as an example of self-administered positive punishment. Heward, Dardig, and Rossett (1979) described the case of a teenage girl who frequently said *don't* instead of *doesn't* with the third person singular (e.g., "She don't like that one.") and used this form of self-administered positive punishment to decrease the rate of her speech error. Each time the girl observed herself saying *don't* when she should have said *doesn't*, she covertly repeated the complete sentence she had just spoken 10 times in a row, using correct grammar. She wore a wrist counter that reminded her to listen to her speech and to keep track of the number of times she employed the positive practice procedure.

Self-Management Analogs of Negative Punishment

Self-administered analogs to negative punishment consist of arranging the loss of reinforcers (response cost) or denying oneself access to reinforcement for a specified period (time-out) contingent on the occurrence of the target behavior. Response cost and time-out contingencies are widely used self-management strategies. The most commonly applied self-administered response cost procedure is paying a small fine each time the target behavior occurs. In one study, smokers reduced their rate of smoking by tearing up a dollar bill each time they lit up (Axelrod, Hall, Weis, & Rohrer, 1971). Response cost procedures have also been used effectively with elementary school children who self-determined the number of tokens they should lose for inappropriate social behavior (Kaufman & O'Leary, 1972) or for poor academic work (Humphrey, Karoly, & Kirschenbaum, 1978).

James (1981) taught an 18-year-old man who had stuttered severely since the age of 6 to use a time-out from speaking procedure. Whenever he observed himself stuttering, the young man immediately stopped talking for at least 2 seconds, after which he could begin speaking again. His rate of dysfluencies decreased markedly. If talking is reinforcing, then this procedure might function as time-out (e.g., not allowing oneself to engage in a preferred activity for a period of time).

Recommendations for Self-Administered Consequences

Persons designing and implementing self-administered consequences should consider the following recommendations.

Select Small, Easy-to-Deliver Consequences

Both rewards and penalties used in self-management programs should be small and easily delivered. A common mistake in designing self-management programs is selecting consequences that are big and grand. Although a person may believe that the promise of a large reward (or threat of a severe aversive event) will motivate her to meet her self-determined performance criteria, large consequences often work against the program's success. Self-selected rewards and punishing consequences should not be costly, elaborate, time-consuming, or too severe. If they are, the person may not be able (in the case of grand rewards) or willing (in the case of severe aversive events) to deliver them immediately and consistently.

In general, it is better to use small consequences that can be obtained immediately and frequently. This is particularly important with consequences intended to function as punishers, which—to be most effective—must be delivered immediately every time the behavior targeted for reduction is emitted.

Set a Meaningful but Easy-to-Meet Criterion for Reinforcement

When designing contingencies involving self-administered consequences, a person should guard against making the same two mistakes practitioners often make when implementing reinforcement contingencies with clients: (1) setting expectations so low that improvement in the current level of performance is not necessary to obtain the self-administered reward, or (2) making the initial performance criterion too high (the more common mistake), thereby effectively programming an extinction contingency that may cause the person to give up on self-management altogether. Keys to the effectiveness of any reinforcement-based intervention are setting an initial criterion that ensures that the person's behavior makes early contact with reinforcement and that continued reinforcement requires improvements over baseline levels. The criterion-setting formulae (Heward, 1980) described in Chapter 11 provide guidelines for making those determinations.

Eliminate "Bootleg Reinforcement"

Bootleg reinforcement—access to the specified reward or to other equally reinforcing items or events without meeting the response requirements of the contingency—is a common downfall of self-management projects. A person with a comfortable supply of smuggled rewards is less likely to work hard to earn response-contingent rewards.

Bootleg reinforcement is common when people use everyday preferred activities and treats as rewards in self-management programs. Although everyday expectations and enjoyments are easy to deliver, a person may find it difficult to withhold things he is used to enjoying on a regular basis. For example, a man used to unwinding at the end of the day by watching *Baseball Tonight* with a beer and a bowl of peanuts may not be consistent in making those daily treats contingent on meeting the requirements of his self-management program.

One method for combating this form of bootleg reinforcement is to retain free access to activities or items routinely enjoyed in the past and make alternative items or activities that are a cut above the usual contingent on meeting performance criteria. For example, each time the man in the previous scenario meets his performance criteria, he could replace his everyday brand of beer with his choice of a specialty beer from a collection reserved in the back of the refrigerator (e.g., Bell's Two Hearted IPA).

If Necessary, Put Someone Else in Control of Delivering Consequences

Most ineffective self-management programs fail not because the controlling behavior is ineffective in controlling the controlled behavior, but because the contingencies that control the controlling behavior are not sufficiently strong. In other words, the person does not emit the controlling behavior consistently enough to realize its effects. What keeps a person from rationalizing that she did most of the behaviors she was supposed to and obtaining a self-determined reward anyway? What keeps a person from failing to deliver a self-determined aversive consequence? Too often, the answer to both questions is—nothing.

A person who really wants to change her behavior but has difficulty following through with delivering self-determined consequences should enlist another person to serve as performance manager. A self-manager can ensure her self-designed consequences will be administered faithfully by creating a contingency in which the consequence for failing to meet the performance criterion is both aversive to her *and* a reinforcer for the person she has asked to implement the consequence. If the first person put in charge of the contingency fails to implement it as planned, the self-manager should find another person who will do the job.

> Christie wanted to walk on her dust-gathering treadmill 20 minutes a day, 6 days a week. She tried her husband as her performance contractor, but he wasn't tough enough; he always felt sorry for her. So she fired him and hired her son; she made his bed every time she failed to do her 20 minutes, and he showed no mercy. (Malott, 2012, p. 132)

Internet-based self-deposit contracts for cigarette smoking reduction or abstinence (Dallery, Raiff, & Grabinski, 2013; Jarvis & Dallery, 2017) or those created on stickK.com ensure that self-determined consequences are carried out. Kanfer (1976) called this kind of self-management *decisional self-control*: The person makes the initial decision to alter her behavior and plans how that will be accomplished, but then turns over the procedure to a second party to avoid the possibility of not emitting the controlling response. Kanfer made a distinction between decisional self-control and protracted self-control, in which a person consistently engages in self-deprivation to effect the desired behavior change. Bellack and Hersen (1977) stated that decisional self-control "is generally considered to be less desirable than protracted self-control, as it does not provide the individual with an enduring skill or resource" (p. 111).

We disagree that a self-management program involving the help of other people is less desirable than one in which the self-manager does everything. First, a self-management program

in which the contingencies are turned over to someone else may be more effective than trying to "go it alone" because the other person is more consistent in applying the consequences. In addition, as a result of experiencing a successful self-management program in which he has chosen the target behavior, set performance criteria, determined a self-monitoring/self-evaluation system, and arranged to have someone else administer self-designed consequences, a person has acquired a considerable repertoire of self-management skills for future use.

Keep It Simple

A person should not create elaborate self-management contingencies if they are not needed. The same general rule that applies to behavior change programs designed on behalf of others—employ the least complicated and intrusive, yet effective, intervention—also applies to self-management programs.

OTHER SELF-MANAGEMENT TACTICS

Other self-management strategies used by behavior analysts are not so easily classified according to the four-term contingency. They include self-instruction, habit reversal, self-directed systematic desensitization, and massed practice.

Self-Instruction

People talk to themselves all the time, offering encouragement (e.g., "I can do this; I've done it before"), congratulations (e.g., "Great shot, Daryl! You just crushed that 5-iron!"), and admonishment (e.g., "Don't say that anymore; you hurt her feelings") for their behavior, as well as specific instructions (e.g., "Pull the bottom rope through the middle"). Such self-statements can function as controlling responses—verbal mediators—that affect the occurrence of other behaviors. **Self-instruction** consists of self-generated verbal responses, covert or overt, that function as response prompts for a desired behavior. As a self-management tactic, self-instructions are often used to guide oneself through a behavior chain or sequence of tasks.

Bornstein and Quevillon (1976) conducted a study frequently cited as evidence of the positive and lasting effects of self-instruction. They taught three hyperactive preschool boys a series of four types of self-instruction designed to keep them on task with classroom activities:

1. Questions about the assigned task (e.g., "What does the teacher want me to do?")
2. Answers to the self-directed questions (e.g., "I'm supposed to copy that picture.")
3. Verbalizations to guide the child through the task at hand (e.g., "OK, first I draw a line through here.")
4. Self-reinforcement (e.g., "I really did that one well.")

During a 2-hr session, the children were taught to use self-instructions incorporating a sequence of training steps originally developed by Meichenbaum and Goodman (1971):

1. The experimenter modeled the task while talking aloud to himself.
2. The child performed the task while the experimenter provided the verbal instructions.
3. The child performed the task while talking aloud to himself with the experimenter whispering the instructions softly.
4. The child performed the task while whispering softly to himself with the experimenter moving his lips but making no sound.
5. The child performed the task while making lip movements but no sound.
6. The child performed the task while guiding his performance with covert instructions. (adapted from p. 117)

During the training session, a variety of classroom tasks were used, ranging from simple motor tasks such as copying lines and figures to more complex tasks such as block design and grouping tasks. The children showed a marked increase in on-task behavior immediately after receiving self-instruction training, and their improved behavior was maintained over a considerable period. The authors suggested that the generalization obtained from the training setting to the classroom could have resulted from telling the children during training to imagine they were working with their teacher, not the experimenter. The authors hypothesized that a behavioral trapping phenomenon (Baer & Wolf, 1970) may have been responsible for the maintenance of on-task behavior; that is, the self-instructions may have initially produced better behavior, which in turn produced teacher attention that maintained the on-task behavior.

Although some studies evaluating self-instruction failed to reproduce Bornstein and Quevillon's (1976) impressive results (e.g. Billings & Wasik, 1985; Friedling & O'Leary, 1979), other studies have yielded generally encouraging results (Burgio, Whitman, & Johnson, 1980; Hughes, 1992; Kosiewicz, Hallahan, Lloyd, & Graves, 1982). Self-instruction training has increased the rate of high school students' initiating conversations with familiar and unfamiliar peers (Hughes, Harmer, Killian, & Niarhos, 1995; see Figure 30.5).

Employees with disabilities have learned to self-manage their work performance by providing their own verbal prompts and self-instructions (Hughes, 1997). For example, Salend, Ellis, and Reynolds (1989) used a self-instruction strategy to teach four adults with severe cognitive disabilities to "talk while you work." Productivity increased dramatically and error rates decreased when the women verbalized to themselves, "Comb up, comb down, comb in bag, bag in box" while packaging combs in plastic bags. Hughes and Rusch (1989) taught two supported employees working at a janitorial supply company how to solve problems by using a self-instruction procedure consisting of four statements:

1. Statement of the problem (e.g., "Tape empty")
2. Statement of the response needed to solve the problem (e.g., "Need more tape")
3. Self-report (e.g., "Fixed it")
4. Self-reinforcement (e.g., "Good")

O'Leary and Dubey (1979) summarized their review of self-instruction training by suggesting four factors that appear to influence its effectiveness with children.

> Self-instructions appear to be effective self-controlling procedures if the children actually implement the instructional procedure, if the children use them [the self-instructions] to influence behavior at which they are skilled, if children have been reinforced for adhering to their self-instructions in the past, and if the focus of the instructions is the behavior most subject to consequences. (p. 451)

Habit Reversal

In his discussion of self-control, Skinner (1953) included "doing something else" as a self-management tactic. In an interesting application of "doing something else," Robin, Schneider, and Dolnick (1976) taught 11 primary-age children with emotional and behavioral disorders to control their aggressive behaviors by using the turtle technique: The children pulled their arms and legs close to their bodies, put their heads down on their desks, relaxed their muscles, and imagined they were turtles. The children were taught to use the turtle response whenever they believed an aggressive exchange with someone else was about to take place, when they were angry with themselves and felt they were about to have a tantrum, or when the teacher or a classmate called out, "Turtle!"

Azrin and Nunn (1973) developed an intervention they called **habit reversal**, in which clients are taught to self-monitor their nervous habits and interrupt the behavior chain as early as possible by engaging in behavior incompatible with the problem behavior (i.e., doing something else). For example, when a nail biter observes herself beginning to bite her fingernails, she might squeeze her hand into a tight fist for 2 or 3 minutes (Azrin, Nunn, & Frantz, 1980) or sit on her hands (Long, Miltenberger, Ellingson, & Ott, 1999). Habit reversal is typically implemented as a treatment package that includes self-awareness training involving response detection and procedures for identifying events that precede and trigger the response and competing response training. Motivation techniques including self-administered consequences, social support systems, and procedures for promoting the generalization and maintenance of treatment gains may also be included. Habit reversal has proven an effective self-management tactic for a wide variety of problem behaviors (Miltenberger, Fuqua, & Woods, 1998).

Mancuso and Miltenberger (2016) conducted a study using habit reversal with six female college students seeking to improve their public speaking skills. The students filled pauses by uttering nonsense syllables such as "uh," "um," or "er"; tongue clicking (pressing the tongue against the roof of the mouth and releasing the pressure to make an audible "tsk" or click sound); and interjecting the word "like." These target behaviors were recorded during 3- to 5-min speeches throughout the study. Each session the researcher randomly selected 2 topics from a list of 25 (e.g., my favorite movie, my first job) and instructed the participant to choose 1. The participant was then given 10 minutes to prepare an outline for a 5-min speech. Participants were allowed to use their outlines during their speeches.

Intervention consisted of one-on-one habit reversal training sessions with the researcher. Response detection began with each participant identifying occurrences of the target behaviors in 3-min video clips of her baseline speeches. The participant then practiced detecting the target behaviors during live speeches by raising her right hand each time she engaged in one of the behaviors and her left hand when becoming aware that she was about to emit one of the behaviors. Response detection training ended when the participant identified 100% of target behavior occurrences in one speech or 85% of the occurrences in two consecutive speeches.

Next, each participant practiced emitting competing responses for each target behavior: for "uh," "um," or "er," a 3-sec silent pause; for tongue clicking, placing the tongue against the inside of the bottom teeth for 3 seconds; for interjecting "like," beginning the sentence again without "like." Competing response training concluded when the participant achieved an 80% reduction in the target behavior from her baseline mean during a speech.

Habit reversal training resulted in an immediate and lasting decrease in filled pauses by the participants (see Figure 29.8). The participants emitted moderate to high rates of filled pauses during baseline (mean across participants of 7.4 responses per minute). Following habit reversal training, the mean rate of filled pauses by the participants was 1.4 responses per minute. Each participant maintained her improved performance in a follow-up speech 2 to 5 weeks later.

Self-Directed Systematic Desensitization

Systematic desensitization is a widely used behavior therapy treatment for anxieties, fears, and phobias that features the self-management strategy of engaging in an alternative behavior (i.e., doing something else). Originally developed by Wolpe (1958, 1973), **systematic desensitization** involves substituting one behavior, generally muscle relaxation, for the unwanted behavior—the fear and anxiety. The client develops a hierarchy of situations from the least to the most fearful and then learns to relax while imagining these anxiety-producing situations, first the least fearful situation, then the next fearful one, and so on. Figure 29.9 shows an anxiety-stimulus hierarchy that a person might develop in attempting to control a fear of cats. When a person is able to go completely through his hierarchy, imagining each scene in detail while maintaining deep relaxation and feeling no anxiety, he begins to expose himself gradually to real-life situations.

Detailed procedures for achieving deep muscle relaxation, constructing and validating a hierarchy of anxiety- or fear-producing situations, and implementing a self-directed systematic desensitization program can be found in Martin and Pear (2015) and Wenrich, Dawley, and General (1976).

Massed Practice

Forcing oneself to perform an undesired behavior again and again, a technique called **massed practice**, will sometimes

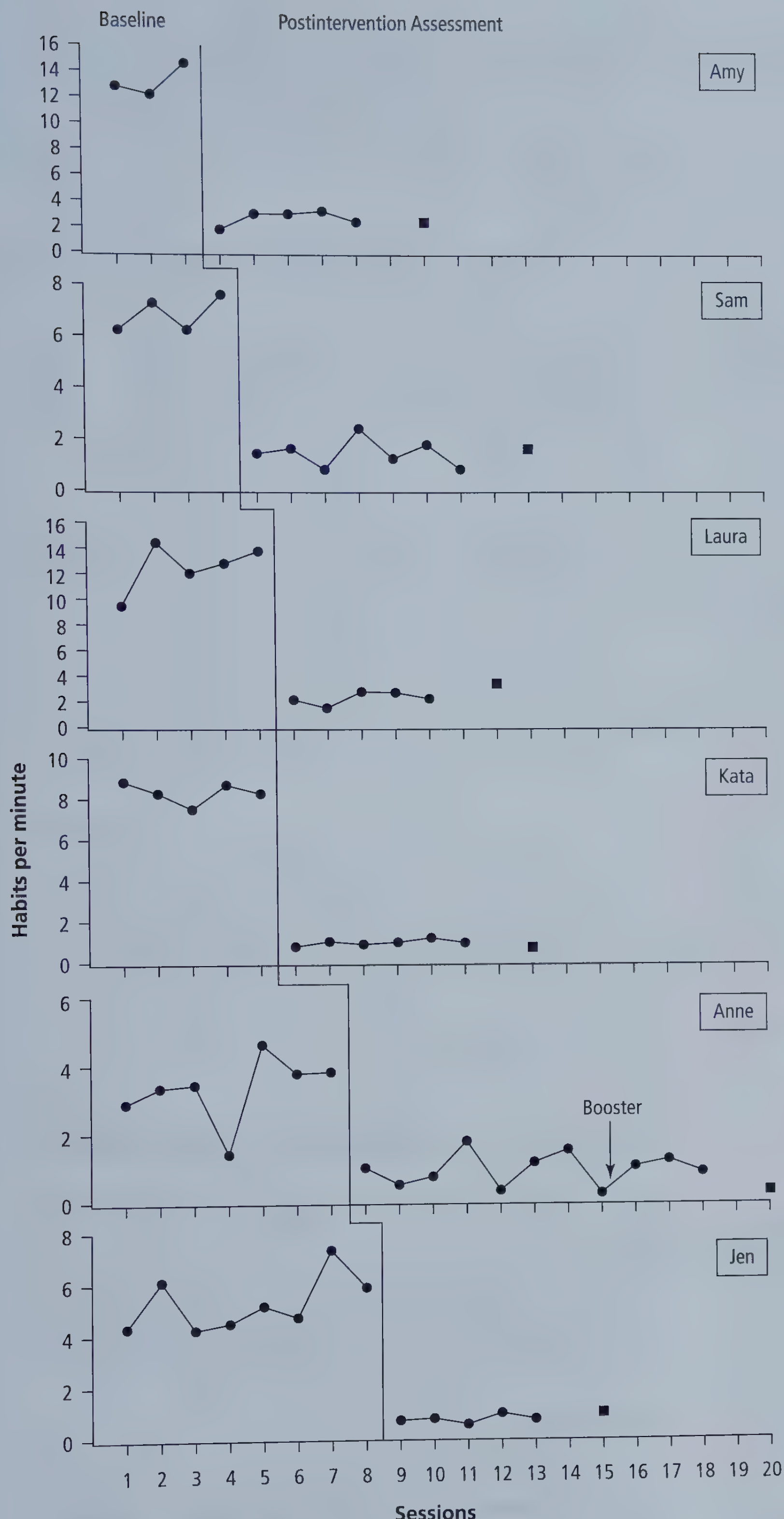


Figure 29.8 Effects of habit reversal training on the rate of filled pauses during speeches by six university students during baseline and after habit reversal training. The downward arrow represents sessions conducted after Anne received a booster session. Square data points represent 2- to 5-week follow-up speeches.

From "Using Habit Reversal to Decrease Filled Pauses in Public Speaking." by C. Mancuso and R. G. Miltenberger, 2016, *Journal of Applied Behavior Analysis*, *49*, p. 191. Copyright 2016 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by Permission.

decrease the future occurrences of the behavior. Wolff (1977) reported an interesting case of this form of treatment by a 20-year-old woman who engaged in a compulsive, ritualized routine of 13 specific security checks every time she entered her apartment (e.g., looking under beds, checking the closets, looking in the kitchen). She began her program by deliberately going through the 13 steps in an exact order and then repeating the complete ritual 4 more times. After doing this for 1 week, she permitted herself to check the apartment if she wanted to but made herself go through the entire routine 5 times whenever she did any checking at all. She soon stopped performing her compulsive checking behavior.

Figure 29.9 Sequence of imaginary scenes concerning fear of cats that could be used for systematic self-desensitization.

Directions

1. You're sitting in a comfortable chair in the safety of your home watching TV.
2. You're watching a commercial for cat food—no cat is visible.
3. The commercial continues and a cat is now eating the food.
4. A man is now petting the cat.
5. A man is holding the cat and fondling it.
6. A woman is holding the cat, and the cat is licking her hands and face.
7. You're looking out the window of your home and you see a cat on the lawn across the street.
8. You're sitting in front of your house, and you see a cat walk by on the sidewalk across the street.
9. You're sitting in your yard, and you see a cat walk by on your sidewalk.
10. A cat walks within 15 feet of you.
11. A friend of yours picks the cat up and plays with it.
12. Your friend is 10 feet away from you, and the cat is licking his face.
13. Your friend comes within 5 feet of you while he's holding the cat.
14. Your friend stands 2 feet away and plays with the cat.
15. Your friend asks if you'd like to pet the cat.
16. Your friend reaches out and offers you the cat.
17. He puts the cat on the ground, and it walks over to you.
18. The cat rubs up against your leg.
19. The cat walks between your legs purring.
20. You reach down and touch the cat.
21. You pet the cat.
22. You pick up the cat and pet it. (p. 71)

SUGGESTIONS FOR EFFECTIVE SELF-MANAGEMENT PROGRAMS

Incorporating the suggestions that follow into the design and implementation of self-management programs should increase the likelihood of success. Although none of these guidelines has been rigorously examined through experimental analyses—research on self-management has a long way to go—each suggestion is consistent with procedures proven effective in other applications of behavior analysis and with "best practices" commonly recommended in the self-management literature (e.g., Malott, 2012; Martin & Pear, 2015; Watson & Tharp, 2014).

1. Specify a goal and define the behavior to be changed.
2. Self-monitor the behavior.
3. Contrive contingencies that will compete with natural contingencies.
4. Go public with your commitment to change your behavior.
5. Get a self-management partner.
6. Continually evaluate your self-management program and redesign it as necessary.

Specify a Goal and Define the Target Behavior

A self-management program begins with identifying a personal goal or objective and the specific behavior changes necessary to accomplish that goal or objective. A person can use most of the questions and issues that practitioners should consider when selecting target behaviors for students or clients (see Chapter 3) to assess the social significance and prioritize the importance of a list of self-determined target behavior changes.

Self-Monitor the Behavior

A person should begin self-monitoring as soon as he has defined the target behavior. Self-monitoring prior to implementing any other form of intervention yields the same benefits as does taking baseline data described in Chapter 7.

1. Self-monitored baseline data provide information about antecedent-behavior-consequent correlations that may be helpful in designing an effective intervention.
2. Self-monitored baseline data provide guidance for setting initial performance criteria for self-administered consequences.

3. Self-monitored baseline data provide an objective basis for evaluating the effects of any subsequent interventions.

Another reason to begin self-monitoring as soon as possible without using additional self-management tactics is that the desired improvement in behavior may be achieved by self-monitoring alone.

Contrive Contingencies That Will Compete with Ineffective Natural Contingencies

When self-monitoring alone does not result in the desired behavior change, the next step is designing a contrived contingency to compete with the natural contingencies. A person who implements a contingency that provides immediate, definite consequences for each occurrence (or perhaps, nonoccurrence) of the target behavior greatly increases the probability of obtaining a previously elusive self-management goal. For example, a smoker who self-records and reports each cigarette smoked to his self-management partner, who provides contingent praise, rewards, and penalties, has arranged immediate, frequent, and more effective consequences for reducing his smoking behavior than those provided by the natural contingency: the threat of emphysema and death by lung cancer in the future, with neither outcome becoming appreciably more immediate or likely from one cigarette to the next.

Go Public

Publicly sharing the intentions of a self-management effort may enhance its effectiveness. When a person shares a goal or makes a prediction to others about her future behavior, she has arranged potential consequences—praise or condemnation—for her success or failure in meeting that goal. A person should state in specific terms what she intends to do and her deadline for completing it. Taking the idea of public commitment a step further, consider the powerful potential of public posting (e.g., an untenured junior faculty member could post a chart of her writing where her department chair or dean could see and comment on it).

Malott (1981) called this the "public spotlight principle of self-management."

> A public statement of goals improves performance. But just how do the social contingencies engaged by public commitment produce change? They increase the rewarding value of success and the aversive value of failure, I presume. But those outcomes are probably too delayed to directly reinforce problem solving. Instead they must be part of the rules the student self-states at crucial times: "If I don't make my goal, I'll look like a fool; but if I do make it, I'll look pretty cool." Such rules then function as cues for immediate self-reinforcement of on-task behavior and self-punishment of off-task behavior. (Volume II, No. 18, p. 5)

Skinner (1953) also theorized about the principles of behavior that operate when a person shares a self-management goal with important others.

> By making it in the presence of people who supply aversive stimulation when a prediction is not fulfilled, we arrange consequences which are likely to strengthen the behavior resolved upon. Only by behaving as predicted can we escape the aversive consequences of breaking our resolution. (p. 237)

Get a Self-Management Partner

Setting up a self-management exchange is a good way to involve another person whose differential feedback about how a self-management project is going can be effective as a behavioral consequence. Two people, each of whom has a long-range goal or a regular series of tasks to do, agree to interact with each other—in person or by phone, text, e-mail, or social media—on a daily or weekly basis, as determined by their schedules, preferences, the target behaviors, and each person's progress. They share their self-monitoring data and exchange verbal praise or admonishments and perhaps even more tangible consequences contingent on performance. Malott (1981) reported a successful self-management exchange in which he and a colleague paid each other \$1 every time either person failed to complete any one of a series of self-determined daily exercise, housekeeping, and writing tasks. Each morning they spoke on the telephone and reported their performance during the previous 24 hours.

To help themselves and each other study for exams and complete research and writing tasks, a group of doctoral students created a self-management exchange they called the Dissertation Club (Ferreri et al., 2006). At weekly meetings, each group member reported data on the "scholarly behaviors" she had targeted (e.g., graphs of words written each day or number of hours studied). Group members provided social support to one another in the form of encouragement to continue working hard and praise for accomplishments, served as behavioral consultants to one another on the design of self-management interventions, and sometimes administered rewards and fines to one another. All six members of the group wrote and successfully defended their dissertations within the time lines they had set.

Continually Evaluate and Redesign Program as Needed

> *The biggest problem with self-management . . . is that the systems are liable to gradually deteriorate . . . So, you must always be on top of it. And you must be prepared with duct tape and bubble gum to stick the system back together.*
>
> —R. W. Malott (2012, p. 31)

Most self-management programs embody a pragmatic, data-based approach to personal problem solving more than rigorous research with an emphasis on experimental analysis and control. Like researchers, however, self-managers should be guided by the data. If the data show the program is not working satisfactorily, the intervention should be redesigned.

People who have learned to define, observe, record, and graph their own behavior can evaluate their self-management efforts with self-experiments. A simple A-B design (Chapter 7) provides a straightforward accounting of the results in a before-and-after fashion and is usually sufficient for self-evaluation.

Experimentally determining functional relations between self-managed interventions and their effects, though valuable, usually takes a back seat to the pragmatic goal of changing one's behavior. However, the changing criterion design (Chapter 9) lends itself nicely not only to the stepwise increments in performance that are often a part of improving personal performance, but also to a clearer demonstration and understanding of the relation between the intervention and changes in the target behavior.

In addition to data-based evaluation, people should judge the social validity of their self-management projects (Wolf, 1978). Practitioners teaching self-management to others can help their clients assess the social validity of their self-management efforts by providing checklists of questions covering topics such as how practical their intervention was, whether they believed their self-management program affected their behavior in any unmeasured ways, and whether they enjoyed the project.

One of the more important aspects of social validity for any behavior change program is the extent to which the results—measured changes in the target behavior—actually make a difference in the lives of the participants. One approach to assessing the social validity of the results of a self-management program is to collect data on what Malott (2012) called *benefit measures*. For example, a person might measure:

- Number of pounds lost as a benefit measure of eating less or exercising more
- Improved lung capacity (peak airflow volume measured as cubic centimeters per second on an inexpensive device available at many pharmacies) as an outcome of reducing the number of cigarettes smoked
- Decrease in time it takes to run a mile as a benefit of number of miles run each day
- Lower resting heart rate and faster recovery time as a benefit of aerobic exercise
- Higher scores on foreign language practice exams as an outcome of studying

In addition to assessing the social validity of self-management interventions, the positive results of benefit measures may serve as consequences that reward and strengthen continued adherence to self-management.

Box 29.2: "Take a Load Off Me" describes a self-management weight loss program that incorporated many of the tactics and suggestions described in this chapter.

BEHAVIOR CHANGES BEHAVIOR

Referring to a book on self-management he authored for a popular audience, Epstein (1997) wrote:

> A young man whose life is in disarray (he smokes, drinks, overeats, loses things, procrastinates, and so on) seeks advice from his parents, teachers, and friends, but no one can help. Then the young man remembers his Uncle Fred (modeled shamelessly after Fred Skinner), whose life always seemed to be in perfect harmony. In a series of visits, Uncle Fred reveals to him the three "secrets" of self-management, all *Ms*: *Modify your environment, monitor your behaviors*, and *make commitments*. Fred also reveals and explains the "self-management principle": *Behavior changes behavior*. After each visit, the young man tries out a new technique, and his life is changed radically for the better. In one scene, he sees a classroom of remarkably creative and insightful children who have been trained in self-management techniques in a public school. It is fiction, of course, but the technology is well established and the possibilities are well within reach. (p. 563, emphasis in original)

Building upon Skinner's (1953) conceptual analysis of self-control, applied behavior analysts have developed numerous self-management tactics and methods for teaching learners with diverse abilities how to apply them. Underlying all of these efforts and findings is the simple, yet profound principle that behavior changes behavior.

BOX 29.2

Take a Load Off Me: A Self-Managed Weight Loss Program

Joe was a 63-year-old whose doctor had recently told him that the 195 pounds on his 5-foot 11-inch frame must be reduced to 175 pounds or he would likely experience serious health problems. Although Joe got regular physical exercise by keeping up the lawn and garden, cutting and hauling firewood, feeding the rabbits, and doing other chores that come with maintaining a farmhouse, his prodigious appetite—long considered one of the "best in the county"—had caught up with him. His doctor's warning and the recent death of a high school classmate scared Joe enough to sit down with his son and plan a self-management weight-loss program.

Joe's program was a package intervention that included antecedent-based tactics, self-monitoring, self-evaluation, contingency contracting, self-selected and self-administered consequences, and contingency management by significant others.

Goal

To reduce the current weight of 195 pounds to 175 pounds at a rate of 1 pound per week.

Behavior Change Needed to Achieve Goal

Reduce eating to a maximum 2100 calories per day.

Rules and Procedures

1. Mount the bathroom scale each morning before dressing or eating and chart the weight on "Joe's Weight" graph taped to the bathroom mirror.
2. Carry a notepad, pencil, and calorie counter in a pocket throughout the day and record the type and amount of *all* food and liquid (other than water) *immediately* after consumption.

3. Before going to bed each night, add the total calories consumed during the day and chart that number on "Joe's Eating" graph taped to the bathroom mirror.
4. *Do not waiver* from Steps 1–3, regardless of weight loss or lack of weight loss.

Immediate Contingencies/Consequences

- The calorie counter, notepad, and pencil in the pocket provide continuously available response prompts.
- Recording all food and drink consumed provides an immediate consequence.
- When you have not eaten any food you thought about eating, put a star on your notepad and give self-praise ("Way to go, Joe!").

Daily Contingencies/Consequences

- If the calories consumed during the day do not exceed the 2100 criterion, put 50 cents in "Joe's Garden Jar."
- If total calories exceed the criterion, remove $1.00 from "Joe's Garden Jar."
- For each 3 days in a row that you meet the calorie criterion, add a bonus of 50 cents to "Joe's Garden Jar."
- Have Helen initial your contract each day that you meet the calorie criterion.

Weekly Contingencies/Consequences

- Each Sunday night, write the calories consumed each day for the previous week on one of the predated, addressed, and stamped postcards. Helen will verify your self-report by initialing the postcard and mailing it to Bill and Jill on Monday.
- Each Monday, if total calories consumed met the criterion for at least 6 of the past 7 days, obtain an item or activity from "Joe's Reward Menu."

Intermediate Contingencies/Consequences

- Eat within the daily calorie limit often enough and there will be sufficient money in "Joe's Garden Jar" to buy seeds and plants for the spring vegetable garden.
- If your weight during the May visit to Ohio represents a loss of at least 1 pound per week, be a guest of honor at a restaurant of your choice.

Long-term Consequences

- Feel better.
- Look better.
- Be healthier.
- Live longer.

Results

Joe often struggled with the rules and procedures, but he stuck with his self-management program and lost 22 pounds (achieving a weight of 173) in 16 weeks. Joe lived 30 active years following his adventure in self-management. He retained his weight at 175 pounds and enjoyed gardening, singing in a barbershop choir, and listening to the Cubs on the radio.

SUMMARY

The "Self" as Behavior Controller

1. We tend to assign causal status to events that immediately precede behavior, and when causal variables are not readily apparent in the immediate, surrounding environment, the tendency to turn to internal causes is strong.

2. Hypothetical constructs such as willpower and drive are explanatory fictions that bring us no closer to understanding the behaviors they claim to explain and lead to circular reasoning.

3. Skinner (1953) conceptualized self-control as a two-response phenomenon: The *controlling response* affects variables in such a way as to change the probability of the other, the *controlled response*.

Self-Management Defined

4. We define self-management as the personal application of behavior change tactics that produces a desired change in behavior.

5. Self-management is a relative concept. A behavior change program may entail a small degree of self-management or be totally conceived, designed, and implemented by the person.

6. Although *self-control* and *self-management* often appear interchangeably in the behavioral literature, we recommend that *self-management* be used in reference to a person acting in some way *in order to* change his subsequent behavior. Three reasons for this recommendation are as follows:
   - Self-control implies that the ultimate control of behavior lies within the person, but the causal factors for "self-control" are to be found in a person's experiences with his environment.
   - Self-control as a causal factor can serve as an explanatory fiction suggesting "a [separate] self inside or [that there is] a self inside controlling external behavior" (Baum, 2017, p. 167).
   - Laypeople and behaviorists alike use self-control to refer to a person's ability to "delay gratification" by responding to achieve a delayed, but larger or higher quality, reward instead of acting to obtain an immediate, less valuable reward.

7. People discount the value of delayed rewards; the greater the delay, the less value or influence the reward has on current behavior. Behavior analysts call this phenomenon delay discounting (or *temporal discounting*).

Applications, Advantages, and Benefits of Self-Management

8. Four uses of self-management are to
   - live a more effective and efficient daily life,
   - break bad habits and acquire good ones,
   - accomplish difficult tasks, and
   - achieve personal lifestyle goals.
9. Advantages and benefits of learning and teaching self-management skills include the following:
   - Self-management can influence behaviors not accessible to external change agents.
   - External change agents often miss important instances of behavior.
   - Self-management can promote the generalization and maintenance of behavior change.
   - A small repertoire of self-management skills can control many behaviors.
   - People with diverse abilities can learn self-management skills.
   - Some people perform better under self-selected tasks and performance criteria.
   - People with good self-management skills contribute to more efficient and effective group environments.
   - Teaching students self-management skills provides meaningful practice for other areas of the curriculum.
   - Self-management is an ultimate goal of education.
   - Self-management benefits society.
   - Self-management helps a person feel free.
   - Self-management feels good.

Antecedent-based Self-Management Tactics

10. Antecedent-based self-management tactics feature the manipulation of events or stimuli antecedent to the target (controlled) behavior, such as the following:
    - Manipulating motivating operations to make a desired (or undesired) behavior more (or less) likely
    - Providing response prompts
    - Performing the initial steps of a behavior chain to ensure being confronted later with a discriminative stimulus that will evoke the desired behavior
    - Removing the materials required for an undesired behavior
    - Limiting an undesired behavior to restricted stimulus conditions
    - Dedicating a specific environment for a desired behavior

Self-Monitoring

11. Self-monitoring is a procedure whereby a person observes and responds to, usually by recording, the behavior he is trying to change.
12. Originally developed as a method of clinical assessment for collecting data on behaviors that only the client could observe, self-monitoring evolved into the most widely used and studied self-management strategy because it often results in desired behavior change.
13. Self-monitoring is often combined with goal setting and self-evaluation. A person using self-evaluation compares her performance with a predetermined goal or standard.
14. Self-monitoring is often part of an intervention that includes reinforcement for meeting either self- or teacher-selected goals.
15. It is difficult to determine exactly how self-monitoring works because the procedure necessarily includes, and is therefore confounded by, private events (covert verbal behavior); it often includes either explicit or implicit contingencies of reinforcement.
16. Children can be taught to self-monitor and self-record their behavior accurately by means of a faded matching technique, in which the child is rewarded initially for producing data that match the teacher's or parent's data. Over time the child is required to match the adult's record less often, eventually monitoring the behavior independently.
17. Accuracy of self-monitoring is neither necessary nor sufficient to achieve improvement in the behavior being monitored.
18. Suggested guidelines for self-monitoring are as follows:
    - Provide materials that make self-monitoring easy.
    - Provide supplementary cues or prompts.
    - Self-monitor the most important dimension of the target behavior.
    - Self-monitor early and often, but do not interrupt the flow of a desired behavior targeted for increase.
    - Reinforce accurate self-monitoring.

Self-Administered Consequences

19. As a technical term, *self-reinforcement* (as also, self-punishment) is a misnomer. Although behavior can be changed by self-administered consequences, the variables influencing the controlling response make such self-management tactics more than a straightforward application of operant reinforcement.
20. Self-administered contingencies analogous to positive and negative reinforcement and positive and negative punishment can be incorporated into self-management programs.
21. When designing self-management programs involving self-administered consequences, a person should:
    - Select small, easy-to-deliver consequences.
    - Set a meaningful but easy-to-meet criterion for reinforcement.
    - Eliminate "bootleg reinforcement."
    - If necessary, put someone else in control of delivery consequences.
    - Use the least complicated and intrusive contingencies that will be effective.

Other Self-Management Tactics

22. Self-instruction (talking to oneself) can function as a controlling response (verbal mediator) that affects the occurrence of other behaviors.

23. Habit reversal is a multiple-component treatment package in which clients are taught to self-monitor their unwanted habits and interrupt the behavior chain as early as possible by engaging in a behavior incompatible with the problem behavior.

24. Systematic desensitization is a behavior therapy treatment for anxieties, fears, and phobias that involves substituting one behavior, generally muscle relaxation, for the unwanted behavior—the fear and anxiety. Self-directed systematic desensitization involves developing a hierarchy of situations from the least to the most fearful and then learning to relax while imagining these anxiety-producing situations, first the least fearful situation, then the next fearful one, and so on.

25. Massed practice, forcing oneself to perform an undesired behavior again and again, can decrease future occurrences of the behavior.

Suggestions for Effective Self-Management Programs

26. Following are six steps in designing and implementing a self-management program:

Step **1:** Specify a goal and define the behavior to be changed.

Step **2:** Self-monitor the behavior.

Step **3:** Create contingencies that will compete with natural contingencies.

Step **4:** Go public with the commitment to change behavior.

Step **5:** Get a self-management partner.

Step **6:** Continually evaluate and redesign the program as needed.

Behavior Changes Behavior

27. The most fundamental principle of self-management is that behavior changes behavior.

KEY TERMS

habit reversal
massed practice
self-control
self-evaluation
self-instruction
self-management
self-monitoring
systematic desensitization

MULTIPLE-CHOICE QUESTIONS

1. The personal application of behavior change tactics that produces a desired change in behavior is:

a. Self-management

b. Managing behaviors

c. Behavior management

d. Behavior modification

Hint: (See "The 'Self' as Controller of Behavior")

2. Self-management can help individuals to:

a. Break bad habits and replace with good ones

b. Accomplish difficult tasks

c. Achieve personal goals

d. All of these

Hint: (See "Applications, Advantages, and Benefits of Self-Management")

3. Limiting an undesired behavior to restructure stimulus conditions is

a. A consequent self-management tactic

b. An antecedent-based self-management tactic

c. Always used when possible to produce an effective self-management program

d. None of these

Hint: (See "Antecedent-Based Self-Management Tactics")

4. Self-monitoring is best defined as:

a. A procedure whereby persons observe their own behavior and record only the occurrence of the target behavior

b. A procedure whereby persons observe the behavior of someone else and systematically record the occurrence or nonoccurrence of the target behavior

c. A procedure whereby persons observe their behavior systematically and record the occurrence or nonoccurrence of the target behavior

d. A procedure whereby persons observe their behavior and someone else systematically records the occurrence or nonoccurrence of the target behavior.

Hint: (See "Self-Monitoring")

5. Suggested guidelines for effective and efficient self-monitoring include:

a. Make materials cumbersome and sophisticated

b. Never provide supplementary cues or prompts; this will make the individual become dependent on the prompts or cues rather than the target behavior

c. Self-monitor the most important dimension of the target behavior

d. Do not reinforce accurate self-monitoring

Hint: (See "Self-Monitoring")

6. Self-administered consequences that increase desired behavior include:
   a. Self-management analogs of positive reinforcement
   b. Self-management analogs of positive punishment
   c. Self-management analogs of negative punishment
   d. None of these
   Hint: (See "Self-Administered Consequences")
7. If you wanted to decrease a desired behavior, what type of self-administered consequence would be most useful?
   a. Self-management analogs of positive reinforcement
   b. Self-management analogs of negative reinforcement
   c. Self-management analogs of positive punishment
   d. None of these
   Hint: (See "Self-Administered Consequences")
8. A type of self-management tactic that typically involves the use of self-awareness training, competing response training, and motivation techniques is:
   a. Self-Instruction
   b. Habit Reversal
   c. Massed Practice
   d. None of these
   Hint: (See "Other Self-Management Tactics")
9. This type of self-management tactic involves self-generated verbal responses, covert or overt, that function as response prompts for a desired behavior.
   a. Self-instruction
   b. Self-directed systematic desensitization
   c. Habit reversal
   d. None of these
   Hint: (See "Other Self-Management Tactics")
10. The first step in conducting an effective self-management program is:
    a. Begin self-monitoring the behavior
    b. Contrive contingencies that will compete with natural contingencies
    c. Continually evaluate your self-management program and redesign it as necessary
    d. Specify a goal and define the behavior to be changed
    Hint: (See "Suggestions for Conducting an Effective Self-Management Program")

ESSAY-TYPE QUESTIONS

1. What is self-management and what are three of its potential advantages or benefits?
   Hint: (See "The 'Self' as a Controller of Behavior" and "Applications, Advantages, and Benefits of Self-Management")
2. What is self-monitoring and what procedures promote its effectiveness?
   Hint: (See "Self-Monitoring")
3. What are self-administered consequences and what is an example of one that would increase a desired or decrease an undesired behavior?
   Hint: (See "Self-Administered Consequences")
4. Name and describe an additional self-management tactic that is discussed within the text.
   Hint: (See "Other Self-Management Tactics")
5. List and discuss the steps for conducting an effective self-management program.
   Hint: (See "Suggestions for Conducting an Effective Self-Management Program")

NOTES

1. Explanatory fictions and circular reasoning are discussed in Chapters 1 and 11.
2. There is also nothing new about self-management. As Epstein (1997) noted, many self-control techniques outlined by Skinner were described by ancient Greek and Roman philosophers and have appeared in the teachings of organized religions for millennia (cf. Bolin & Goldberg, 1979; Shimmel, 1977, 1979).
3. For a variety of conceptual analyses of self-control/self-management from a behavioral perspective, see Brigham (1983), Catania (1975, 2013), Goldiamond (1976), Hughes and Lloyd (1993), Kanfer and Karoly (1972), Malott (1984, 1989, 2005a, b), Moore (2015), Rachlin (1995, 2016), and Watson and Tharp (2014).
4. In early biological theory, *homunculus* was a fully formed human being thought to exist inside an egg or spermatozoon.
5. Generalized behavior change is a defining goal of applied behavior analysis and the focus of Chapter 30.
6. Motivating operations are described in detail in Chapter 16.
7. The critical importance of immediacy in reinforcement is discussed in Chapter 11.

PART 12

Promoting Generalized Behavior Change

It is one thing to implement interventions that change behavior; it is quite another to extend those behavior changes through time, across settings, or to new behaviors. Practitioners face no more challenging or important task than that of designing, implementing, and evaluating interventions that produce behavior changes that continue after the intervention is terminated, appear in relevant environments beyond the training setting, and engender related behaviors that were not taught directly. Chapter 30 defines the major types of generalized behavior change and describes strategies and tactics applied behavior analysts use to achieve them.

CHAPTER 30

Generalization and Maintenance of Behavior Change

LEARNING OBJECTIVES

- Define and provide examples of the different basic forms of generalized behavior change.
- Define and provide examples of the different undesirable types of generalized behavior change.
- Discuss planning techniques for generalized behavior change.
- List and discuss strategies and tactics for promoting generalized behavior change.
- Discuss how to modify and terminate successful interventions.
- Name and discuss the guiding principles for promoting generalized outcomes.

Sherry's teacher implemented an intervention that taught her to complete each part of multiple-part, in-school assignments before submitting them and beginning another activity. Now, 3 weeks after the program ended, most of the work Sherry submits as "finished" is incomplete and her stick-with-a-task-until-it's-finished behavior is as poor as it was before the intervention began.

Ricardo has just begun his first competitive job working as a copy machine operator in a downtown business office. In spite of his long history of distractibility and poor endurance, Ricardo had learned to work independently for several hours at a time in the copy room at the vocational training center. His employer, however, is complaining that Ricardo frequently stops working after a few minutes to seek attention from others. Ricardo may soon lose his job.

Brian is a 10-year-old boy diagnosed with autism. In an effort to meet an objective on his individualized education program that targets functional language and communication skills, Brian's teacher taught him to say, "Hello, how are you?" as a greeting. Now, whenever Brian meets anyone, he invariably responds with, "Hello, how are you?" Brian's parents are concerned that their son's language seems stilted and parrot-like.

The mere passage of time resulted in Sherry losing her ability to complete assignments. A change of scenery threw Ricardo off his game; the excellent work habits he acquired at the vocational training center were absent at the community job site. Although Brian used his new greeting skill over time and with new people, its repetitive and restricted form did not serve him well.

Each of these three situations illustrates a common type of teaching failure. The most socially significant behavior changes last over time, are used in all relevant settings and situations, and are accompanied by other functionally related responses. The student who learns to count money and make change in the classroom today must be able to count and make change at the convenience store tomorrow and at the supermarket next month. The beginning writer who has been taught to write a few good sentences in school must be able to write many more meaningful sentences when writing notes or letters to family or friends.

Applied behavior analysts face no more challenging or important task than that of designing, implementing, and evaluating interventions that produce generalized behavior change. This chapter defines three major types of generalized behavior change and describes the strategies and tactics researchers and practitioners use most often to promote them.

GENERALIZED BEHAVIOR CHANGE: DEFINITIONS AND KEY CONCEPTS

Baer, Wolf, and Risley (1968) included **generalized behavior change** as one of seven defining characteristics of applied behavior analysis.

> A behavior change may be said to have generality if it proves durable over time, if it appears in a wide variety of possible environments, or if it spreads to a wide variety of related behaviors. (p. 96)

In their seminal review paper, "An Implicit Technology of Generalization," Stokes and Baer (1977) also stressed those three facets of generalized behavior change—across time, settings, and behaviors—when they defined *generalization* as

> the occurrence of relevant behavior under different, nontraining conditions (i.e., across subjects, settings,

> people, behaviors, and/or time) without the scheduling of the same events in those conditions. Thus, generalization may be claimed when no extratraining manipulations are needed for extratraining changes; or may be claimed when some extra manipulations are necessary, but their cost is clearly less than that of the direct intervention. Generalization will not be claimed when similar events are necessary for similar effects across conditions. (p. 350)

In other words, when a learner emits a trained behavior in places or at times without having to be retrained completely in those places or times, or emits functionally related behaviors that were not taught directly, generalized behavior change is evident. The following sections define and provide examples of these three basic forms of generalized behavior change: response maintenance, setting/situation generalization, and response generalization.

Response Maintenance

We define **response maintenance** as the extent to which a learner continues to perform a target behavior after a portion or all of the intervention responsible for the behavior's initial appearance in the learner's repertoire has been terminated.[1] For example:

- Sayaka was having difficulty identifying the lowest common denominator (LCD) when adding and subtracting fractions. Her teacher had Sayaka write the steps for finding the LCD on an index card and told her to refer to the card when needed. Sayaka began using the LCD cue card, and the accuracy of her math assignments improved. After using the cue card for a week, Sayaka said she no longer needed it and returned it to her teacher. The next day Sayaka correctly computes the LCD for every problem on a quiz on adding and subtracting fractions.
- On Loraine's first day on the job with a residential landscaping company, a coworker taught her how to use a long-handled tool to extract dandelions. Without further instruction, Loraine continues to use the tool correctly a month later.
- When Derek was in the seventh grade, one of his teachers taught him to record his assignments and keep materials for each class in separate folders. Now a college sophomore, Derek continues to apply those organizational skills to his academic work.

Response maintenance was evident in Sayaka's performance on a math quiz 1 day after the cue card intervention ended, in Loraine's effective use of the weeding tool 1 month after training, and in Derek's continued use of the organizational skills he had learned years earlier. How long a newly learned behavior must be maintained depends on how long the learner's environment presents opportunities or requirements to perform the behavior. If covertly reciting a phone number 3 times immediately after hearing it enables a person without a cell phone to remember the number long enough to tap it in on a borrowed phone a few minutes later, sufficient response maintenance has been achieved. Other behaviors, such as literacy, self-care, and social skills, must be maintained in a person's repertoire for a lifetime.

Meaningful response maintenance may be obtained with some components of the initial intervention in place. A study by Sprague and Horner (1984), in which six students with moderate to severe cognitive disabilities were taught to use vending machines, provides a good example of this point. The students were given cue cards to aid them in operating vending machines without assistance. The cue cards had food and drink logos on one side and pictures of quarters paired with prices on the other. The cards were used during instruction and generalization probes, but also were kept by the students at the end of the program. Five of the six students still carried a cue card and used vending machines independently 18 months after the study had ended.

Setting/Situation Generalization

Setting/situation generalization is the extent to which a learner emits a target behavior in a setting or stimulus situation that differs from the instructional setting in any meaningful way. For example:

- While waiting for his new motorized wheelchair to arrive from the factory, Chaz used a computer simulation program and a joystick to learn how to operate his soon-to-arrive chair. When the new chair arrived, Chaz grabbed the joystick and immediately began zipping up and down the hall and spinning perfectly executed donuts.
- Loraine was taught to pull weeds from flowerbeds and mulched areas. Although never having been instructed to do so, Loraine began removing weeds from lawns she crossed on her way to the flowerbeds.
- After Brandy's teacher taught her to read 10 different C-V-C-E words (e.g., *bike, cute, made*), Brandy could read other C-V-C-E words for which she had not received instruction (e.g., *bite, cake, mute*).

A study by van den Pol and colleagues (1981) provides an excellent example of setting/situation generalization. They taught three young adults with multiple disabilities to eat independently in fast-food restaurants. All three students had previously eaten in restaurants, but could not order or pay for a meal without assistance. The researchers began by constructing a task analysis of the steps required to order, pay for, and eat a meal appropriately in a fast-food restaurant. Instruction took place in the students' classroom and consisted of role playing each of the steps during simulated customer–cashier interactions and responding to questions about photographic slides showing customers at a fast-food restaurant performing the various steps in the sequence. The 22 steps in the task analysis were divided into four major components: locating, ordering, paying, and eating and exiting. After a student had mastered the steps in each component in the classroom, he was given "a randomly determined number of bills equaling two to five dollars and instructed to go eat lunch" at a local restaurant (p. 64). Observers stationed inside the restaurant recorded each student's performance of each step in the task analysis. The results of these generalization probes, which

were also conducted before training (baseline) and after training (follow-up), are shown in Figure 30.1. In addition to assessing the degree of generalization from the classroom, which was based on the specific McDonald's restaurant used for most of the probes, the researchers conducted follow-up probes in a Burger King restaurant (also a measure of maintenance).

This study exemplifies the pragmatic approach used by most applied behavior analysts to assess and promote generalized behavior change. The setting in which generalized responding is desired can contain one or more components of the intervention that was implemented in the instructional environment, but not all of the components. If the complete intervention program is required to produce behavior change in a novel environment, then no setting/situation generalization can be claimed. However, if some component or components of the training program result in meaningful behavior change in a generalization setting, then setting/situation generalization can be claimed, provided it can be shown that the component or components used in the generalization setting were insufficient to produce the behavior change alone in the training environment.

For example, van den Pol and colleagues taught Student 3, who was deaf, how to use a prosthetic ordering device in the classroom. The device, a plastic laminated sheet of cardboard with a wax pencil, had preprinted questions (e.g., "How much is . . . ?"), generic item names (e.g., large hamburger), and spaces where the cashier could write responses. Simply giving the student some money and the prosthetic ordering card would not have enabled him to order, purchase, and eat a meal independently. However, after classroom instruction that included guided practice, role playing, social reinforcement ("Good job! You remembered to ask for your change" [p. 64]), corrective feedback, and review sessions with the prosthetic ordering card produced the desired behaviors in the instructional setting. Student 3 was able to order, pay for, and eat meals in a restaurant, aided only by the card.

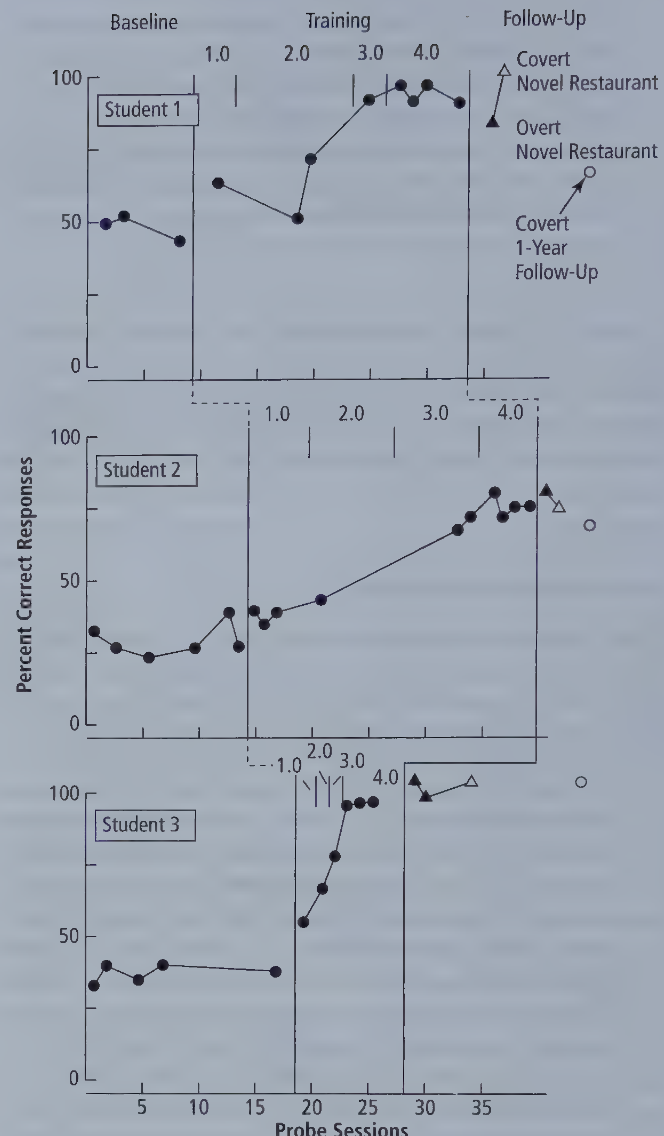


Figure 30.1 Percentage of steps necessary to order a meal at a fast-food restaurant correctly performed by three students with disabilities before, during, and after instruction in the classroom. During follow-up, the closed triangles represent probes conducted at a Burger King restaurant using typical observation procedures, open triangles represent Burger King probes during which students did not know they were being observed, and open circles represent covert probes conducted in a different McDonald's 1 year after training.

From "Teaching the Handicapped to Eat in Public Places: Acquisition, Generalization and Maintenance of Restaurant Skills" by R. A. van den Pol, B. A. Iwata, M. T. Ivanic, T. J. Page, N. A. Neef, and F. P. Whitley, 1981, *Journal of Applied Behavior Analysis, 14,* p. 66. Reproduced with permission of John Wiley & Sons, Inc.

Distinguishing Between Instructional and Generalization Settings

Instructional setting denotes the environment where teaching occurs, including all planned or unplanned aspects of the environment that may influence the learner's acquisition and generalization of the target behavior.[2] Planned elements are the stimuli and events the teacher has programmed to achieve initial behavior change and promote generalization. For example, planned elements for a math lesson would include the specific problems and solution steps to be presented and the format and sequencing of those items. Unplanned aspects of the instructional setting are factors the teacher is not aware of or has not considered that might affect the acquisition and generalization of the target behavior. For example, the phrase *how much* in a word problem may acquire stimulus control over a student's use of addition, even when the correct solution to the problem requires a different arithmetic operation. Or, perhaps a student always uses subtraction for the first problem on each page of word problems because a subtraction problem has always been presented first during instruction.

A **generalization setting** is any place or stimulus situation that differs from the instructional setting in some meaningful way and in which performance of the target behavior is desired. There are multiple generalization settings for many target behaviors. The student who learns to solve addition and subtraction word problems in the classroom must be able to solve similar problems at home, at the store, and in the dugout with teammates.

Examples of instructional and generalization settings for six target behaviors are shown in Figure 30.2. When a person uses a skill in an environment physically removed from the setting where she learned it—as with behaviors 1 through 3 in Figure 30.2—it is easy to understand that event as an example of generalization across settings. However, many important generalized outcomes occur across more subtle differences between the instructional setting and generalization setting. A generalization setting does not have to be some*where* different from the physical location where instruction is provided. Students often receive instruction in the same place where they will need to maintain and generalize what they have learned. In other words, the instructional setting and generalization setting can, and often do, share the same physical location (as with behaviors 4 through 6 in Figure 30.2).

Distinguishing Between Setting/Situation Generalization and Response Maintenance

Because any measure of setting/situation generalization is conducted after some instruction has taken place, it might be argued that setting/situation generalization and response maintenance are the same, or are inseparable phenomena at least. Most measures of setting/situation generalization provide information on response maintenance, and vice versa. For example, the post-training generalization probes conducted by van den Pol and colleagues (1981) at the Burger King restaurant and at the second McDonald's provided data on setting/situation generalization (i.e., to novel restaurants) and on response maintenance of up to 1 year. However, a functional distinction exists between setting/situation generalization and response maintenance, with each outcome presenting a somewhat different set of challenges for programming and ensuring enduring behavior change. When a behavior change produced in the classroom or clinic is not observed in the generalization environment, a lack of setting/situation generalization is evident. When a behavior change produced in the classroom or clinic has occurred at least once in the generalization setting and then ceases to occur, a lack of response maintenance is evident.

An experiment by Koegel and Rincover (1977) illustrates the functional difference between setting/situation generalization and response maintenance. Participants were three young boys with autism; each was mute, echolalic, or displayed no appropriate contextual speech. One-to-one instructional sessions were conducted in a small room with the trainer and child seated across from each other at a table. Each child was taught to imitate a series of responses (e.g., the trainer said, "Touch your [nose, ear]" or "Do this" and [raised his arm, clapped his hands]). Each 40-min session consisted of blocks of 10 training trials in the instructional setting alternated with blocks of 10 trials conducted by an unfamiliar adult standing outside, surrounded by trees. All correct responses in the instructional setting were followed by candy and social praise. During generalization trials, the children received the same instructions and model prompts as in the classroom, but no reinforcement or other consequences were provided for correct responses.

Figure 30.3 shows the percentage of trials in which each child responded correctly in the instructional setting and in the generalization setting. All three children learned to respond to the imitative models in the instructional setting. At the end of the experiment, all three children showed 0% correct responding in the generalization setting, but for different reasons. Child 1 and Child 3 began emitting correct responses in the generalization setting as their performances improved in the instructional setting; however, their generalized responding was not maintained (most likely the result of the extinction condition in effect in the generalization setting). The imitative responding acquired by Child 2 in the instructional setting never generalized to the outside setting. Therefore, the 0% correct responding at the experiment's conclusion confirms a lack of response maintenance for Child 1 and Child 3, but for Child 2 it represents a failure of setting generalization.

Figure 30.2 Examples of a possible instructional setting and a generalization setting for various target behaviors.

Instructional Setting	Generalization Setting
1. Raising hand when *special education teacher* asks a question in the *resource room.*	1. Raising hand when *general education teacher* asks a question in the *regular classroom.*
2. Practicing conversational skills with *speech therapist at school.*	2. Talking with *peers in town.*
3. Passing basketball during a *team scrimmage on home court.*	3. Passing basketball during a *game on the opponent's court.*
4. Answering *addition problems in vertical format* at desk at school.	4. Answering *addition problems in horizontal format* at desk at school.
5. Solving *word problems with no distracter* numbers on homework assignment.	5. Solving *word problems* with *distracter numbers* on homework assignment.
6. Operating package sealer at community job site *in presence of supervisor.*	6. Operating package sealer at community job site *in absence of supervisor.*

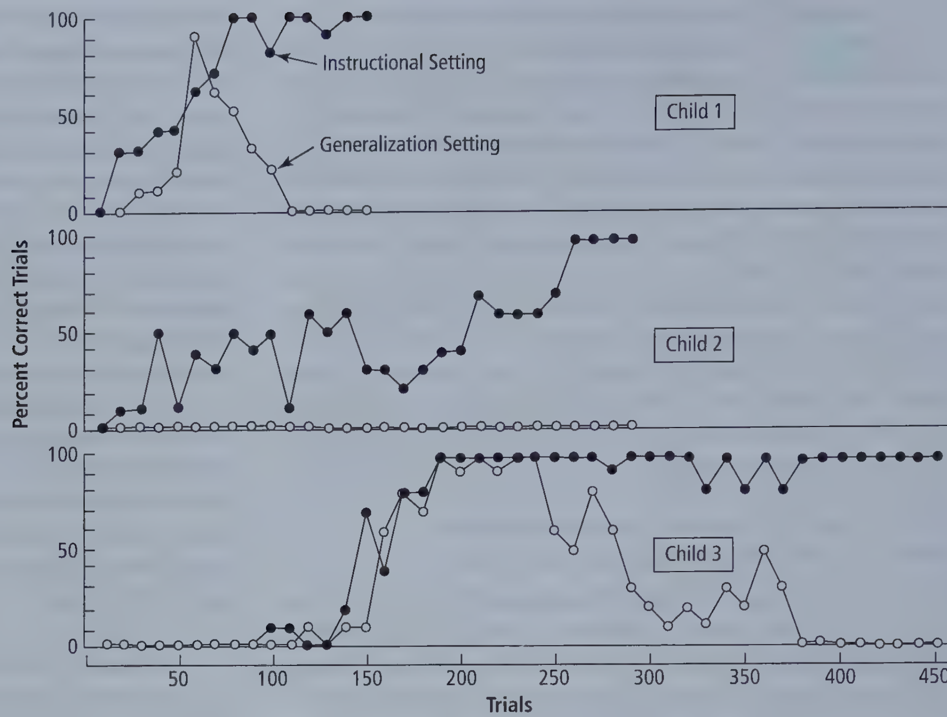


Figure 30.3 Correct responding by three children on alternating blocks of 10 trials in the instructional setting and 10 trials in the generalization setting.

From "Research on the Differences Between Generalization and Maintenance in Extra-Therapy Responding" by R. L. Koegel and A. Rincover, 1977, *Journal of Applied Behavior Analysis, 10,* p. 4. Reproduced with permission of John Wiley & Sons, Inc.

Response Generalization

Response generalization is the extent to which a learner emits untrained responses that are functionally equivalent to the trained target behavior. In other words, in response generalization useful variations of the target behavior appear for which no instructional contingencies have been applied. For example:

- Traci wanted to earn some extra money by helping her older brother with his lawn mowing business. Her brother taught Traci to walk the mower up and down parallel rows that progressed from one side of a lawn to the other. Traci discovered that she could mow some lawns just as quickly by first cutting around the perimeter and then walking the mower in a concentric pattern toward the center of the lawn.
- Loraine was taught to remove weeds with a long weed-removal tool. Although never taught or asked to do so, sometimes Loraine removes weeds with a hand trowel or with her bare hands.
- Michael's mother taught him how to take phone messages by using the pencil and notepaper next to the phone to write the caller's name, phone number, and message. One day, Michael's mother came home and saw her son's tape recorder next to the phone. She pushed the play button and heard Michael's voice say, "Grandma called. She wants to know what you'd like her to cook for dinner Wednesday. Mr. Stone called. His number is 555-1234. He said your car is ready to be picked up."

The study by Goetz and Baer (1973) described in Chapter 8 of three preschool girls' block building provides a good example of response generalization. During baseline, the teacher sat by each girl as she played with the blocks, watching closely but quietly, and displaying neither enthusiasm nor criticism for any particular use of the blocks. During the next condition, each time the child placed or rearranged the blocks to create a form that had not appeared previously in that session's constructions, the teacher commented with enthusiasm and interest (e.g., "Oh, that's very nice—that's different!"). Another condition followed in which each repeated construction of a form within the session was praised (e.g., "How nice—another arch!"). The study ended with a phase in which descriptive praise was again contingent on the construction of different block forms. All three children constructed more new forms with the blocks when form diversity was praised than under baseline or the praise-for-the-same-forms condition (see Figure 8.7).

Even though specific responses produced reinforcement (i.e., the actual block forms that preceded each instance of teacher praise), other responses sharing that functional characteristic (i.e., being different from block forms constructed previously by the child) increased in frequency as a function of the teacher's praise. As a result, during reinforcement for different forms, the children constructed new forms with the blocks even though each new form itself had never before appeared and therefore could not have been reinforced previously. Reinforcing a few members of the response class of new forms increased the frequency of other members of the same response class.

Generalized Behavior Change: A Relative and Intermingled Outcome

Generalized behavior change is a relative concept; it exists along a continuum. At one end of the continuum is a minimum amount of generalized behavior change—the learner uses the new skill in a limited range of nontraining settings and situations where contrived response prompts and/or consequences still might need to be applied. At the other end of the continuum lies maximum generalized behavior change; that is, after all components of an

intervention have been terminated, the learner emits the newly acquired target behavior and several functionally related behaviors not observed previously at every appropriate opportunity, in all relevant settings, and performs them indefinitely and independently.

We have presented the three primary types of generalized behavior change individually to isolate their defining features, but they often overlap and occur in combination. Although it is possible to obtain response maintenance without generalization across settings/situations or behaviors (i.e., the target behavior continues to occur in the same setting in which it was trained after the training contingencies have been terminated), all meaningful measures of setting generalization also entail some degree of response maintenance. And it is common for all three forms of generalized behavior change to be represented in the same instance. For example, during a relatively quiet shift at the widget factory on Monday, Joyce's supervisor taught her to obtain assistance by calling out, "Ms. Johnson, I need some help." Later that week (response maintenance) when it was very noisy on the factory floor (setting/situation generalization), Joyce signaled her supervisor by waving her hand back and forth (response generalization).

Generalized Behavior Change Is Not Always Desirable

It is hard to imagine any behavior important enough to target for systematic instruction for which response maintenance would be undesirable.[3] However, unwanted setting/situation generalization and response generalization occur often, and practitioners should design intervention plans to prevent or minimize such unwanted outcomes. Undesirable setting/situation generalization takes two common forms: overgeneralization and faulty stimulus control.

Overgeneralization, a nontechnical but effectively descriptive term, refers to an outcome in which the behavior has come under the control of an overly broad stimulus class. That is, the learner emits the target behavior in the presence of stimuli that, although similar in some way to the instructional examples or situation, are inappropriate occasions for the behavior. For example, a student learns to spell *division, mission*, and *fusion* with the grapheme *–sion*. When asked to spell *fraction*, the student writes *f-r-a-c-s-i-o-n*.

With *faulty stimulus control*, the target behavior comes under the restricted control of an irrelevant antecedent stimulus. For example, after learning to solve word problems such as "Ling has 3 books. Nikola has 5 books. How many books do they have in total?" by adding the numerals in the problem, the student adds the numerals in any problem containing "in total" (e.g., "Raluca has 3 candies. Corinne and Raluca have 8 candies in total. How many candies does Corrine have?").[4]

Undesired response generalization occurs when a learner's untrained but functionally equivalent responses result in poor performance or undesirable outcomes. For example, although Jack's supervisor at the widget factory taught him to operate the drill press with two hands because that is the safest method, sometimes Jack operates the press with one hand. One-handed and two-handed responses are functionally equivalent because both topographies produce widgets, but one-handed responses compromise Jack's health and the factory's safety record. Or, perhaps some of her brother's customers do not like how their lawns look after Traci has mowed them in concentric rectangles.

Box 30.1: "Perspectives on the Sometimes Confusing and Misleading Terminology of Generalization" discusses the varied terms applied behavior analysts use to describe these outcomes.

Other Types of Generalized Outcomes

Other types of generalized outcomes that do not fit easily into categories of response maintenance, setting/situation generalization, and response generalization have been reported in the behavior analysis literature. Complex components of a person's repertoire sometimes appear quickly with little or no apparent direct conditioning. A variety of such emergent performances are described as *stimulus equivalence* relations (Sidman, 1994, 2000), *recombinative generalization* (Axe & Sainato, 2010; Goldstein, 1983), *bidirectional naming* (Horne & Lowe, 1996; Miguel, 2016), and *relational frame theory* (Hayes, Barnes-Holmes, & Roche, 2001) (see chapters 18, 19, and 20). Another type of rapid learning that appears to be a generalized outcome of other events has been called *contingency adduction*, a process whereby behaviors initially established under one set of contingency conditions are recruited and recombined by a different set of contingencies and take on a new function in a person's repertoire (Andronis, Layng, & Goldiamond, 1997; Layng, Twyman, & Stikeleather 2004).

Sometimes an intervention applied to one or more people results in behavior changes in other people who were not directly treated by the contingencies. **Generalization across subjects** refers to changes in the behavior of people not directly treated by an intervention as a function of treatment contingencies applied to other people. This phenomenon, which has been described with a variety of related or synonymous terms—*vicarious reinforcement* (Bandura, 1971; Kazdin, 1973), *ripple effect* (Kounin, 1970), and *spillover effect* (Strain, Shores, & Kerr, 1976)—provides another dimension for assessing the generalization of treatment effects. For example, Fantuzzo and Clement (1981) examined the degree to which behavior changes would generalize from one child who received teacher-administered or self-administered token reinforcement during a math activity to a peer seated next to the child.

Drabman, Hammer, and Rosenbaum (1979) combined four basic types of generalized treatment effects—(a) across time (i.e., response maintenance), (b) across settings (i.e., setting/situation generalization), (c) across behaviors (i.e., response generalization), and (d) across subjects—into a conceptual framework they called the *generalization map*. By viewing each type of generalized outcome as dichotomous (i.e., either present or absent) and by combining all possible permutations of the four categories, Drabman and colleagues arrived at 16 categories of generalized behavior change ranging from maintenance (Class 1) to subject-behavior-setting-time generalization (Class 16). Class 1 generalization is evident if the target behavior of the target subject(s) continues in the treatment setting after any "experiment-controlled contingencies" have been discontinued. Class 16 generalization, the "ultimate form" of generalization, is evidenced by "a change in a nontarget subject's nontarget behavior which endures in a different setting after the contingencies have been withdrawn in the treatment setting" (p. 213).

BOX 30.1

Perspectives on the Sometimes Confusing and Misleading Terminology of Generalization

Applied behavior analysts use a variety of terms to describe behavior changes that appear as adjuncts or by-products of direct intervention. The overlapping and multiple meanings of these terms can lead to confusion and misunderstanding. For example, *maintenance*, the most frequently used term for a behavior change that persists after the intervention that brought it about has been withdrawn or terminated, is also the most common name for an experimental condition in which treatment has been discontinued or partially withdrawn. Applied behavior analysts should distinguish between *response maintenance* as a measure of behavior (i.e., a dependent variable) and *maintenance* as an experimental condition (i.e., an independent variable). Other terms found in the behavior analysis literature for continued responding after programmed contingencies have ended include *durability, behavioral persistence, resurgence*, and (often incorrectly) *resistance to extinction*.

Response maintenance can be measured under extinction conditions, in which case the relative frequency of continued responding is described correctly as *resistance to extinction* (see Chapter 24). However, using *resistance to extinction* to describe response maintenance in most applied situations is incorrect because reinforcement of some kind typically follows some occurrences of the target behavior in the post-treatment environment.

Terms used in the applied behavior analysis literature for behavior changes that occur in nontraining settings or stimulus conditions include *stimulus generalization, setting generalization, transfer of training*, or simply, *generalization*. It is technically incorrect to use *stimulus generalization* to refer to the generalized behavior change achieved by many applied interventions. *Stimulus generalization* is a technical term referring to a specific behavioral process: a response that has been reinforced in the presence of a given stimulus occurring more often in the presence of different but similar stimuli under extinction conditions (Guttman & Kalish, 1956; see Chapter 17). Use of the term should be restricted to that phenomenon.

The occurrence of behaviors that have not been trained directly may be referred to as *collateral* or *side effects, response variability, induction*, or *concomitant behavior change*. To further complicate matters, *generalization* is often used as a catchall term to refer to all three types of generalized behavior change.

Johnston (1979) discussed some problems caused by using *generalization* to describe any desirable behavior change in a generalization setting.

> This kind of usage is misleading in that it suggests that a single phenomenon is at work when actually a number of different phenomena need to be described, explained, and controlled. . . . Carefully designing procedures to optimize the contributions of stimulus and response generalization would hardly exhaust our repertoire of tactics for getting the subject to behave in a desirable way in non-instructional settings. Our successes will be more frequent when we realize that maximizing behavioral influence in such settings requires careful consideration of *all* behavioral principles and processes. (pp. 1–2)

Inconsistent use of the "terminology of generalization" can lead researchers and practitioners to incorrect assumptions and conclusions regarding the principles and processes responsible for the presence or absence of generalized outcomes. Nevertheless, applied behavior analysts will probably continue to use *generalization* as a dual-purpose term, referring sometimes to types of behavior change and sometimes to behavioral principles and processes that produce such changes. Stokes and Baer (1977) clearly indicated their awareness of the differences in definitions.

> The notion of generalization developed here is an essentially pragmatic one; it does not closely follow the traditional conceptualizations (Keller & Schoenfeld, 1950; Skinner, 1953). In many ways, this discussion will sidestep much of the controversy concerning terminology. (p. 350)

While discussing the use of naturally existing contingencies of reinforcement to maintain and extend programmed behavior changes, Baer (1999) explained his preference for using the term *generalization*:

> It is the best of the techniques described here and, interestingly, it does not deserve the textbook definition of "generalization." It is a reinforcement technique, and the textbook definition of generalization refers to unreinforced behavior changes resulting from other directly reinforced behavior changes. . . . [But] we are dealing with the pragmatic use of the word *generalization*, not the textbook meaning. We reinforce each other for using the word pragmatically, and it serves us well enough so far, so we shall probably maintain this imprecise usage. (p. 30, emphasis in original)

In an effort to promote the precise use of the technical terminology of behavior analysis and as a reminder that the phenomena of interest are usually products of multiple principles and procedures, we use terms that focus on the type of generalized behavior change rather than the variables that bring it about.

Although Drabman and colleagues recognized that "the classifications may prove arbitrary" (p. 204), they provided objectively stated rules for determining whether a given event fits the requirements of each of their 16 classifications. Regardless of whether generalized behavior change consists of such distinctly separate and wide-ranging phenomena as detailed by Drabman and colleagues, their generalization map provided an objective framework with which researchers can assess the extended effects of behavioral interventions (Allen, Tarnowski, Simonian, Elliot, & Drabman, 1991). For example, Stevenson and Fantuzzo (1984) measured 15 of the 16 generalization map categories in a study of the effects of teaching a fifth-grade boy to use self-management techniques. They not only measured the effects of the intervention on the target behavior (math performance) in the instructional setting (school), but also assessed effects on the student's math behavior at home, disruptive behavior at home and at school, both behaviors for a nontreated peer in both settings, and maintenance of all of the above.

PLANNING FOR GENERALIZED BEHAVIOR CHANGE

> *In general, generalization should be programmed, rather than expected or lamented.*
>
> —Baer, Wolf, and Risley (1968, p. 97)

In their review of 270 published studies relevant to generalized behavior change, Stokes and Baer (1977) concluded that practitioners should always "assume that generalization does not occur except through some form of programming . . . and act as if there were no such animal as 'free' generalization—as if generalization never occurs 'naturally,' but always requires programming" (p. 365). Of course, generalization of some type and degree usually occurs, whether or not it is planned. Unplanned and unprogrammed generalization may be sufficient, but often it is not, particularly for many learners served by applied behavior analysts (e.g., children and adults with learning problems and developmental disabilities). And if left unchecked, unplanned generalized outcomes may be undesirable outcomes.

Achieving optimal generalized outcomes requires thoughtful, systematic planning. This planning begins with two major steps: (1) selecting target behaviors that will meet natural contingencies of reinforcement and (2) specifying all desired variations of the target behavior and the settings/situations in which those behaviors should (and should not) occur after instruction has ended.

Select Target Behaviors that will Meet Naturally Existing Contingencies of Reinforcement

> *The everyday environment is full of steady, dependable, hardworking sources of reinforcement for almost all of the behaviors that seem natural to us. That is why they seem natural.*
>
> —Donald M. Baer (1999, p. 15)

Numerous criteria have been suggested for determining whether a proposed teaching objective is relevant or functional for the learner. Age appropriateness and the degree to which a skill represents normalization are often cited as important criteria when selecting target behaviors for students with disabilities (e.g., Brown, McDonnell, & Snell, 2016). These criteria were discussed in Chapter 3, along with numerous other issues that should be considered when selecting and prioritizing target behaviors. In the end, however, there is just one ultimate criterion of functionality: *A behavior is functional only to the extent that it produces reinforcement for the learner.* This criterion holds no matter how important the behavior may be to the person's health or welfare, or no matter how much teachers, family, friends, or the learner herself considers the behavior to be desirable. We repeat: A behavior is not functional if it does not produce reinforcement for the learner. We can say this another way: Behaviors that are not followed by reinforcers on at least some occasions will not be maintained.

Ayllon and Azrin (1968) recognized this fundamental truth when they recommended that when selecting target behaviors practitioners follow the *relevance-of-behavior rule*: Choose only those behaviors to change that will produce reinforcers in the learner's postintervention environment. Baer (1999) believed so strongly in the importance of this criterion that he recommended practitioners heed a similar rule:

> *A good rule is to not make any deliberate behavior changes that will not meet natural communities of reinforcement.* Breaking this rule commits you to maintain and extend the behavior changes that you want, by yourself, indefinitely. If you break this rule, do so knowingly. Be sure that you are willing and able to do what will be necessary. (p. 16, emphasis in original)

Promoting for the generalization and maintenance of any behavior change for which a natural contingency of reinforcement exists, no matter the specific tactics employed, consists of getting the learner to emit the behavior in the generalization setting just often enough to contact the contingency. Generalization and maintenance of the behavior from that point forward, while not ensured, is a very good bet. For example, after receiving some basic instruction on how to operate the steering wheel, gas pedal, and brakes on a car, the naturally existing reinforcement and punishment contingencies involving moving automobiles and the road will select and maintain effective steering, acceleration, and braking. Few drivers need booster training sessions on the basic operation of the steering wheel, gas pedal, and brakes.

We define a **naturally existing contingency** as any contingency of reinforcement (or punishment) that operates independent of the behavior analyst's or practitioner's efforts. Naturally existing contingencies include those that operate without social mediation (e.g., walking fast on an icy sidewalk is often punished by a slip and fall) and socially mediated contingencies contrived and implemented by other people in the generalization setting. From the perspective of a special educator who is teaching a set of targeted social and academic skills to students for whom the general education classroom represents the generalization setting, a token economy operated by the general education classroom teacher is an example of the latter type of naturally existing contingency. Although the general education teacher contrived the token economy, it represents a set of naturally existing contingencies because it already operates in the generalization setting.

We define a **contrived contingency** as any contingency of reinforcement (or punishment) designed and implemented by a behavior analyst or practitioner to achieve the acquisition, maintenance, and/or generalization of a targeted behavior change. The token economy in the previous example is a contrived contingency from the perspective of the general education teacher who designed and implemented it.

Practitioners are often charged with the difficult task of teaching important skills for which there are no dependable naturally existing contingencies of reinforcement. In such cases, practitioners should realize and plan for the fact that the generalization and maintenance of target behaviors will have to be supported, perhaps indefinitely, with contrived contingencies.

Specify All Desired Variations of the Behavior and the Settings/Situations Where Those Behaviors Should (and Should Not) Occur

This stage of planning for generalized outcomes includes identifying all the desired behavior changes that need to be made and all the environments and stimulus conditions in which the learner should emit the target behavior(s) after direct training has ceased (Baer, 1999). For some target behaviors, the most important stimulus control for each response variation is clearly defined (e.g., reading C-V-C-E words) and restricted in number (e.g., solving single-digit multiplication facts). For many important target behaviors, however, the learner is likely to encounter a multitude of settings and stimulus conditions in which the behavior, in a variety of response forms, is desired. Only by considering these possibilities prior to instruction can the behavior analyst design an intervention with the best chance of preparing the learner for them.

In one sense, this component of planning for generalized outcomes is similar to preparing a student for a future test without knowing the content or format of all the questions that will be on the test. The stimulus conditions and contingencies of reinforcement in the generalization setting(s) will provide that test to the learner. Planning involves trying to determine what the final exam will cover (type and form of questions), whether any trick questions will appear (e.g., confusing stimuli that might evoke the target response when it should not occur), and whether the learner will need to use his new knowledge or skill in different ways (response generalization).

List All the Behaviors that Need to be Changed

Listing all behaviors that require change is not an easy task, but a necessary one to obtain a complete picture of the teaching task ahead. For example, if the target behavior is teaching Brian, the young boy with autism, to greet people, he should learn a variety of greetings in addition to "Hello, how are you?" Brian may also need to learn behaviors to initiate and participate in conversations, such as responding to questions, taking turns, staying on topic, and so forth. Further, he may need to be taught when and to whom he should introduce himself. Only by having a complete list of all the desired forms of the behavior can the practitioner make meaningful decisions about which behaviors to teach directly and which to leave to generalization.

The practitioner should determine whether and to what extent response generalization is desirable for all of the behavior changes listed, and then make a prioritized list of the variations of the target behavior she would like to see as generalized outcomes.

List All the Settings and Situations in Which the Target Behavior Should Occur

A list should be made of all the desired settings and situations in which the learner will emit the target behavior if optimal generalization is achieved. Will Brian need to introduce himself and talk with children his own age, to adults, to males and females? Will he need to talk with others at home, at school, in the lunchroom, on the playground? Will he be confronted with situations that may appear to be appropriate opportunities to converse but are not (e.g., an unknown adult approaches and offers candy) and for which an alternative response is needed (e.g., walking away, seeking out a known adult). (This kind of analysis often adds extra behaviors to the list of skills to be taught.)

When all of the possible situations and settings have been identified, they should be prioritized according to their importance and the client's likelihood of encountering them. Further analysis of the prioritized environments should then be conducted. What discriminative stimuli usually set the occasion for the target behavior in these various settings and situations? What schedules of reinforcement for the target behavior are typical in these nontraining environments? What kinds of reinforcers are likely to be contingent on the emission of the target behavior in each of the settings? Only when the behavior analyst has answered all of these questions, if not by objective observation, then at least by considered estimation, can she begin to have a full picture of the teaching task ahead.

Is All This Preintervention Planning Worth It?

Obtaining the information just described requires considerable time and effort. Given limited resources, why not design an intervention and immediately begin trying to change the target behavior? It is true that many behavior changes do show generalization, even though the extension of the trained behavior across time, settings, and other behaviors was unplanned and unprogrammed. When target behaviors have been chosen that are truly functional for the subject, and when those behaviors have been brought to a high level of proficiency under discriminative stimuli relevant to generalization settings, the chances of generalization are good. But what constitutes a high level of proficiency for certain behaviors in various settings? What are all of the relevant discriminative stimuli in all of the relevant settings? What are all the relevant settings?

Without a systematic plan, a practitioner will usually be ignorant of the answers to these vital questions. Few behaviors important enough to target have such limited needs for generalized outcomes that the answers to such questions are obvious. Just a cursory consideration of the behaviors, settings, and people related to a child introducing himself revealed numerous factors that may need to be incorporated into an instructional plan. A more thorough analysis would produce many more. In

fact, a complete analysis will invariably reveal more behaviors to be taught than time or resources would ever allow. And Brian—the 10-year-old who is learning to greet people and introduce himself—in all likelihood needs to learn many other skills also, such as self-help, academic, and recreation and leisure skills, to name just a few. Why, then, create all the lists in the first place when everything cannot be taught anyway? Why not just train and hope?[5]

Baer (1999) described six possible benefits of listing all the forms of behavior change and all the situations in which these behavior changes should occur.

1. You now see the full scope of the problem ahead of you, and thus see the corresponding scope that your teaching program needs to have.
2. If you teach less than the full scope of the problem, you do so by choice rather than by forgetting that some forms of the behavior could be important, or that there were some other situations in which the behavior change should or should not occur.
3. If less than a complete teaching program results in less than a complete set of behavior changes, you will not be surprised.
4. You can decide to teach less than there is to learn, perhaps because that is all that is practical or possible for you to do.
5. You can decide what is most important to teach. You can also decide to teach the behavior in a way that encourages the indirect development of some of the other forms of the desired behavior, as well as the indirect occurrence of the behavior in some other desired situations, that you will not or cannot teach directly.
6. But if you choose the option discussed in number 5 above, rather than the complete program implicit in number 1, you will do so knowing that the desired outcome would have been more certain had you taught every desirable behavior change directly. The best that you can do is to encourage the behavior changes that you do not cause directly. So, you will have chosen the option in number 5 either of necessity or else as a well-considered gamble after a thoughtful consideration of possibilities, costs, and benefits. (pp. 10–11)

After determining which behaviors to teach directly and in which situations and settings to teach those behaviors, the behavior analyst is ready to consider strategies and tactics for promoting generalized outcomes.

STRATEGIES AND TACTICS FOR PROMOTING GENERALIZED BEHAVIOR CHANGE

Various conceptual systems and methodological taxonomies for promoting generalized behavior change have been proposed (e.g., Horner, Dunlap, & Koegel, 1988; Osnes & Lieblein, 2003; Stokes & Baer, 1977; Stokes & Osnes, 1989). The conceptual scheme presented here is informed by the work of those authors and others, and by our experiences designing, implementing, and evaluating methods for promoting generalized outcomes and in teaching practitioners to use them. Although numerous methods and techniques have been demonstrated and given a variety of names, most tactics that effectively promote generalized behavior change can be classified under five strategic approaches:

- Teach the full range of relevant stimulus conditions and response requirements.
- Make the instructional setting similar to the generalization setting.
- Maximize the target behavior's contact with reinforcement in the generalization setting.
- Mediate generalization.
- Train to generalize.

In the following sections we describe and provide examples of 13 tactics used by applied behavior analysts to accomplish these five strategies (see Figure 30.4). Although each tactic is described individually, most interventions to promote generalized behavior change combine these tactics (e.g., Bord, Sidener, Reeve, & Disener, 2017; Grossi, Kimball, & Heward, 1994; Hughes, Harmer, Killina, & Niarhos, 1995; Trask-Tyler, Grossi, & Heward, 1994).

Teach the Full Range of Relevant Stimulus Conditions and Response Requirements

> *The most common mistake that teachers make, when they want to establish a generalized behavior change, is to teach one good example of it and expect the student to generalize from that example.*
>
> —Donald M. Baer (1999, p. 18)

Most important behaviors must be performed in various ways across a wide range of stimulus conditions. Consider a person skilled in reading, math, conversing with others, and cooking. That person can read thousands of different words; add, subtract, multiply, and divide any combination of numbers; make a multitude of relevant and appropriate comments when talking with others; and measure, combine, and prepare numerous ingredients in hundreds of recipes. Helping learners achieve such wide-ranging performances presents an enormous challenge to the practitioner.

One approach to this challenge would be to teach every desired form of a target behavior in every setting/situation in which the learner may need that behavior in the future. Although this approach would eliminate the need to program for response generalization and setting/situation generalization (response maintenance would remain an issue), pursuing it is seldom possible and never practical. A teacher cannot provide direct instruction on every printed word a student may encounter, or teach a student every measuring, pouring, stirring, and sautéing movement needed to prepare every dish he may want to make later on. Even for skill areas for which it would be *possible* to teach every possible example—instruction *could* be provided on all 900 single-digit-times-two-digit multiplication problems—to do so would be impractical for many reasons, not the least of which is that the student needs to learn not only many other types of math problems but also skills in other curriculum areas.

Figure 30.4 Strategies and tactics for promoting generalized behavior change.

Teach the Full-Range of Relevant Stimulus Conditions and Response Requirements
1. Teach enough stimulus examples.
2. Teach enough response examples.

Make the Instructional Setting Similar to the Generalization Setting
3. Program common stimuli.
4. Teach loosely.

Maximize Contact with Reinforcement in the Generalization Setting
5. Teach the target behavior to levels of performance required by naturally existing contingencies of reinforcement.
6. Program indiscriminable contingencies.
7. Set behavior traps.
8. Ask people in the generalization setting to reinforce the target behavior.
9. Teach the learner to recruit reinforcement.

Mediate Generalization
10. Contrive a mediating stimulus.
11. Teach self-management skills.

Train to Generalize
12. Reinforce response variability.
13. Instruct the learner to generalize.

A strategy called **teach enough examples** consists of teaching the student to respond to a subset of all of the possible stimulus and response examples and then probing the student's performance on untrained examples.[6] For example, the generalization of a student's ability to solve two-digit-minus-two-digit arithmetic problems with regrouping can be assessed by asking the student to solve several problems of the same type for which no instruction or guided practice has been provided. If the results of this **generalization probe** show that the student responds correctly to untaught examples, then instruction can be halted on this class of problems. If the student performs poorly on the generalization probe, the practitioner teaches additional examples before again assessing the student's performance on a new set of untaught examples. This cycle of teaching new examples and probing with untaught examples continues until the learner consistently responds correctly to untrained examples representing the full range of stimulus conditions and response requirements found in the generalization setting.

Teach Enough Stimulus Examples

The tactic for promoting setting/situation generalization called *teach enough stimulus examples* involves teaching the learner to respond correctly to multiple examples of antecedent stimulus conditions and probing for generalization to untaught stimulus examples. A different stimulus example is incorporated into the teaching program each time a change is made in any dimension of the instructional item itself or the environmental context in which the item is taught. Examples of four dimensions by which different instructional examples can be identified and programmed are the following:

- The specific *item* taught (e.g., multiplication facts: 7×2, 4×5; letter sounds: *a, t*)
- The *stimulus context* in which the item is taught (e.g., multiplication facts presented in vertical format, in horizontal format, in word problems; saying the sound of *t* when it appears at the beginning and end of words: *t*ab, ba*t*)
- The *setting* where instruction occurs (e.g., large-group instruction at school, collaborative learning group, home)
- The *person* doing the teaching (e.g., classroom teacher, peer, parent)

As a general rule, the more examples used during instruction, the more likely the learner will respond correctly to untrained examples or situations. The number of examples that must be taught before sufficient generalization occurs will vary as a function of factors such as the complexity of the target behavior being taught, the teaching procedures employed, the student's opportunities to emit the target behavior under the various conditions, the naturally existing contingencies of reinforcement, and the learner's history of reinforcement for generalized responding.

Sometimes teaching as few as two examples will produce significant generalization to untaught examples. Stokes, Baer, and Jackson (1974) taught a greeting response to four children with severe cognitive disabilities who seldom acknowledged or greeted other people. The senior author, working as a dormitory assistant, used unconditioned reinforcers (potato chips and M&M'S) and praise to shape the greeting response (at least two back-and-forth waves of a raised hand). Then this initial trainer maintained the newly learned hand wave by contriving three to six contacts per day with each of the children in various settings (e.g., playroom, corridor, dormitory, courtyard). Throughout the study, as many as 23 different staff members systematically approached the children during different times of the day in different settings and recorded whether the children greeted them with a hand waving response. If a child greeted a prober with the waving response, the prober responded with "Hello, (name)." Approximately 20 such generalization probes were conducted each day with each child.

Immediately after learning the greeting response with just one trainer, one of the children (Kerry) showed good

setting/situation generalization by using it appropriately in most of her contacts with other staff members. However, the other three children failed to greet staff members most of the time, even though they continued to greet the original trainer on virtually every occasion. A second staff member then began to reinforce and maintain the greeting responses of these three children. As a result of adding the second trainer, the children's greeting behavior showed widespread generalization to the other staff members. Stokes and colleagues' (1974) study is important for at least two reasons. First, it demonstrated an effective method for continual assessment of setting/situation generalization across numerous examples (in this case, people). Second, the study showed that widespread generalization may result from teaching just two examples.

Teach Enough Response Examples

Instruction that provides practice with a variety of response topographies helps ensure the acquisition of desired response forms and promotes response generalization in the form of untrained topographies. Often called **multiple-exemplar training**, this tactic typically incorporates both stimulus and response variations.[7] Multiple-exemplar training has promoted the generalization of a wide range of behaviors by diverse learners: auditory-visual discrimination tasks (Carr, 2003), empathy skills (Sivaraman, 2017), and sharing (Marzullo-Kerth, Reeve, Reeve, & Townsend, 2011) by children with autism; food preparation (Trask-Tyler, Grossi, & Heward, 1994), domestic skills (Neef, Lensbower, Hockersmith, DePalma, & Gray, 1990), vocational skills (Horner, Eberhard, & Sheehan, 1986), and requests for assistance (Chadsey-Rusch, Drasgow, Reinoehl, Halle, & Collet-Klingenberg, 1993) by children and young adults with disabilities; and identifying definitions of behavior analysis terms by university students (Meindl, Ivy, Miller, Neef, & Williamson, 2013).[8]

Four female high school students with moderate cognitive disabilities participated in a study by Hughes and colleagues (1995) that assessed the effects of an intervention they called *multiple-exemplar self-instructional training* on the acquisition and generalization of the students' conversational interactions with peers. The students were recommended for the study because they initiated conversations and responded to peers' efforts to talk with them at "low or nonexistent rates" and seldom maintained eye contact. One of the students, Tanya, had recently been refused a job at a restaurant because of her "reticence and lack of eye contact during her job interview" (p. 202).

A key element of Hughes and colleagues' intervention was practicing a wide variety of conversation starters and statements with different peer teachers. Ten volunteer peer teachers recruited from general education classrooms helped teach conversation skills to the participants. The peer teachers were male and female, ranged in grade level from freshmen to seniors, and represented African American, Asian American, and Euro-American ethnic groups. Instead of learning a few scripted conversation openers, the participants practiced using multiple examples of conversation starters selected from a pooled list of conversation openers used by general education students. Additionally, participants were encouraged to develop individual adaptations of statements, which further promoted response generalization by increasing the number and range of conversational statements that were likely to be used in subsequent conversations.

Before, during, and after multiple-exemplar training, generalization probes were conducted of each participant's use of self-instructions, eye contact, and initiating and responding to conversation partners. The 23 to 32 different students who served as conversation partners for each participant represented the full range of student characteristics in the school population (e.g., gender, age, ethnicity, students with and without disabilities) and included students who were known and unknown to the participants prior to the study. The rate of conversation initiations by all four participants increased during multiple-exemplar training to levels approximating those of general education students and was maintained at those rates after the intervention was terminated completely (see Figure 30.5).

General Case Analysis

Teaching a learner to respond correctly to multiple examples does not guarantee generalized responding to untaught examples. To achieve optimal generalization and discrimination, the behavior analyst must pay close attention to the specific examples used during instruction; not just any examples will do. Maximally effective instructional design requires selecting teaching examples that represent the full range of stimulus situations and response requirements in the natural environment.[9] **General case analysis** (also called *general case strategy*) is a systematic method for selecting teaching examples that represent the full range of stimulus variations and response requirements in the generalization setting (Albin & Horner, 1988; Becker & Engelmann, 1978).

A series of studies by Horner and colleagues demonstrated the importance of teaching examples that systematically sample the range of stimulus variations and response requirements the learner will encounter in the generalization setting (e.g., Horner et al., 1986; Horner & McDonald, 1982; Horner, Williams, & Steveley, 1987). In a classic example of this line of research, Sprague and Horner (1984) evaluated the effects of general case instruction on the generalized use of vending machines by six high school students with moderate to severe cognitive disabilities. The dependent variable was the number of vending machines each student operated correctly during generalization probes of 10 different machines located within the community. For a probe trial to be scored as correct, a student had to perform correctly a chain of five responses (i.e., insert the proper number of coins, activate the machine for the desired item, and so on). The researchers selected the 10 vending machines used to assess generalization because each student's performance on those machines would serve as an index of his performance "across all vending machines dispensing food and beverage items costing between $.20 and $.75 in Eugene, Oregon" (p. 274). None of the vending machines used in the generalization probes was identical to any of the vending machines used during instruction.

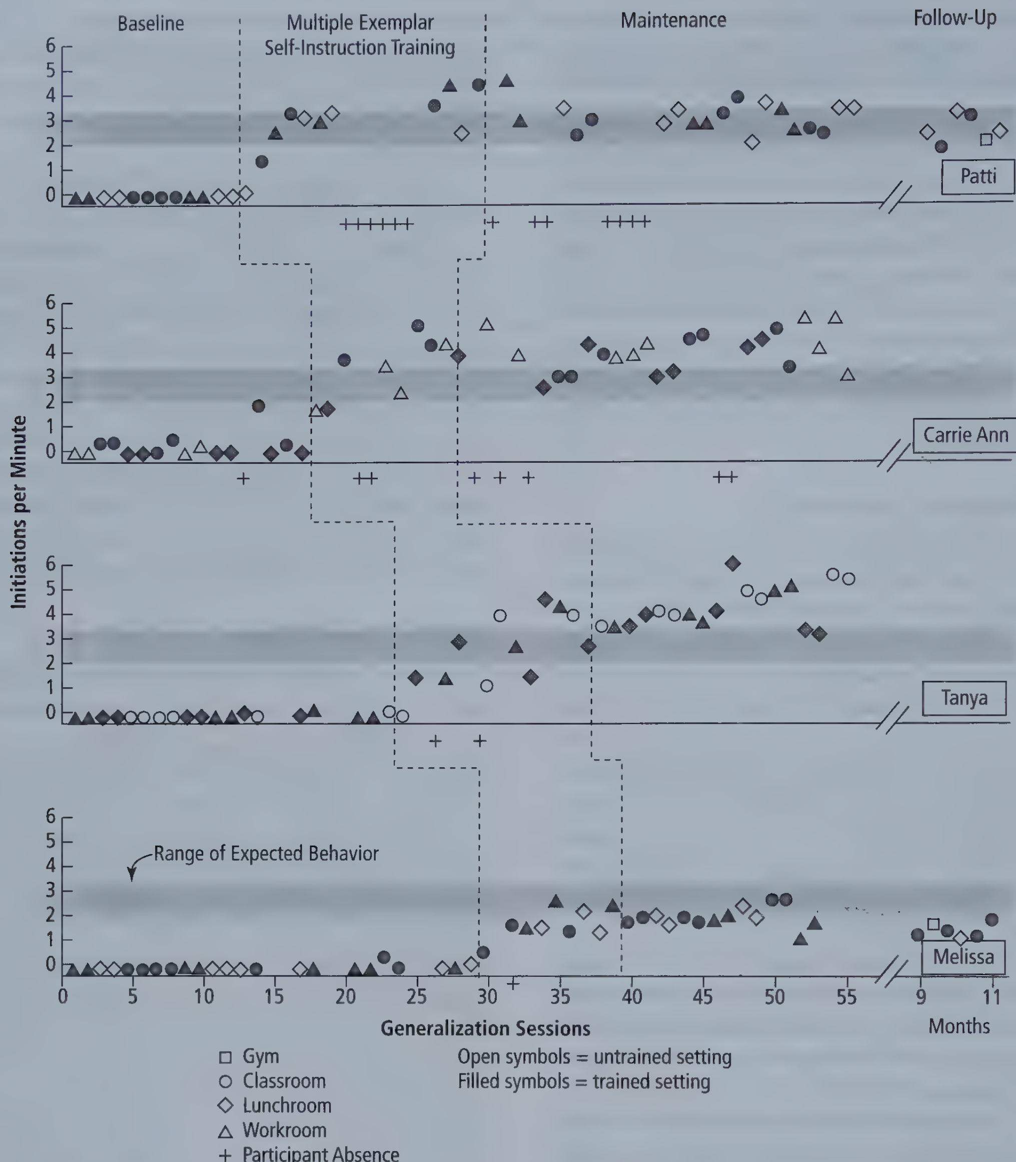


Figure 30.5 Conversation initiations per minute by four high school students with disabilities to conversation partners with and without disabilities during generalization sessions. The shaded bands represent typical performance by general education students.

From "The Effects of Multiple-Exemplar Training on High-School Students' Generalized Conversational Interactions" by C. Hughes, M. L. Harmer, D. J. Killina, and F. Niarhos, 1995. *Journal of Applied Behavior Analysis, 28,* p. 210. Reproduced with permission of John Wiley & Sons, Inc.

After a single-baseline probe verified each student's inability to use the 10 vending machines in the community, a condition the researchers called "single-instance instruction" began. Under this condition, each student received individual training on a single vending machine located in the school until he used the machine independently for three consecutive correct trials on each of 2 consecutive days. Even though each student had learned to operate the training machine without errors, the generalization probe following single-instance instruction revealed little or no success with the vending machines in the community (see Probe Session 2 in Figure 30.6). The continued poor performance by Students 2, 3, 5, and 6 on generalization probes that followed additional instruction with the single-instance training machine shows that "overlearning" on a single example does not necessarily aid generalization. Further evidence of the limited generalization obtained from single-instance instruction is the fact that seven of the eight total probe trials performed correctly by all students after single-instance instruction were on Probe Machine 1, the vending machine that most closely resembled the machine on which the students had been trained.

Next, multiple-instance training was implemented with Students 4, 5, and 6. The teaching procedures and performance criteria for multiple-instance training were the same as those used in the single-instance condition, except that each student received instruction until he reached criterion on three new machines. Sprague and Horner (1984) purposely selected vending machines to use in the multiple-instance instruction that

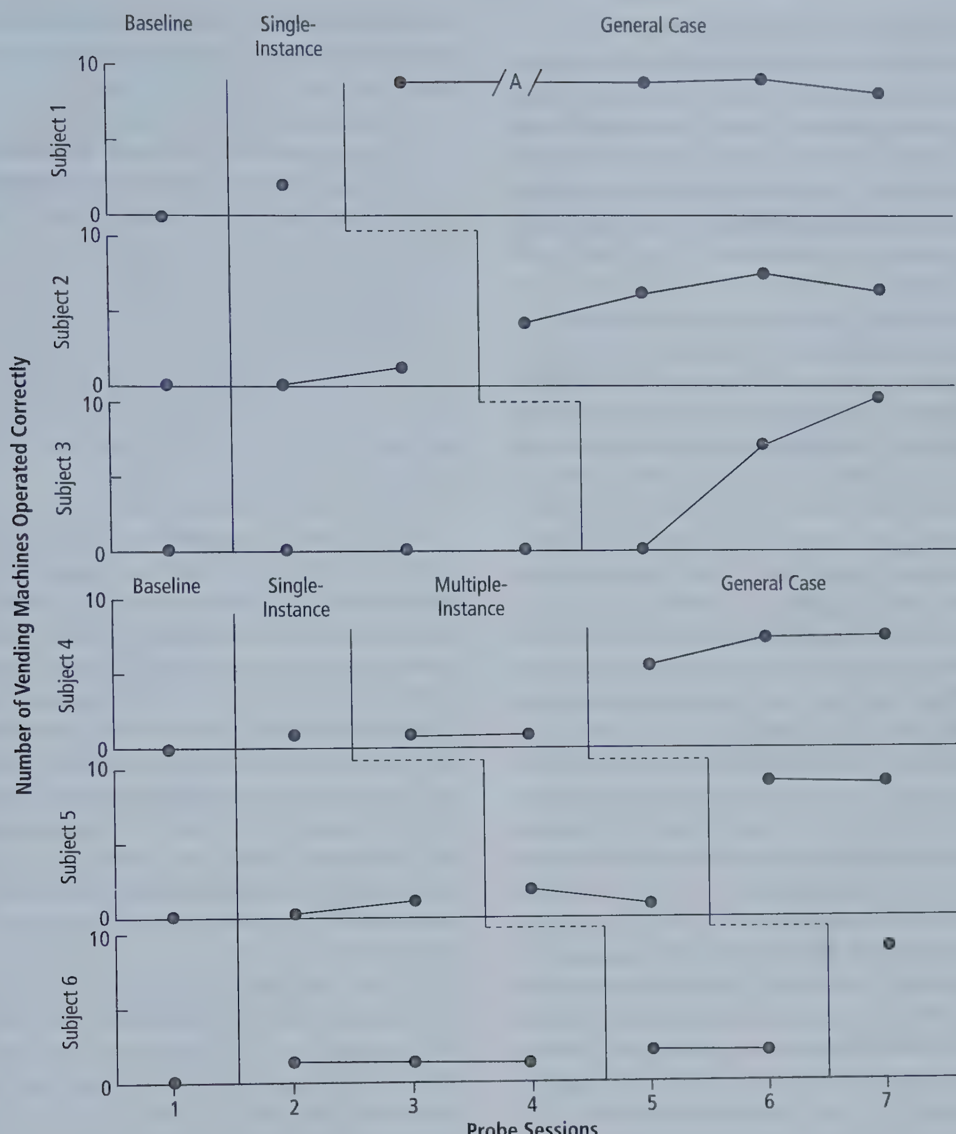


Figure 30.6 Number of nontrained vending machines operated correctly by students across phases and probe sessions.

From "The Effects of Single Instance, Multiple Instance, and General Case Training on Generalized Vending Machine Use by Moderately and Severely Handicapped Students," by J. R. Sprague and R. H. Horner, 1984, *Journal of Applied Behavior Analysis, 17,* p. 276. Reproduced with permission of John Wiley & Sons, Inc.

were similar to one another and that did not sample the range of stimulus variations and response requirements that defined the vending machines in the community. After reaching training criterion on three additional machines, Students 4, 5, and 6 were still unable to operate the machines in the community. During the six probe sessions that followed multiple-instance instruction, these students correctly completed only 9 of the 60 total trials.

The researchers then introduced general case instruction in multiple-baseline fashion across students. This condition was the same as multiple-instance instruction, except that the three different vending machines used in general case training, when combined with the single-instance machine, provided the students with practice across the total range of stimulus conditions and response variations found in vending machines in the community. None of the training machines, however, was exactly the same as any machine used in the generalization probes. After reaching training criterion on the general case machines, all six students showed substantial improvements in their performance on the 10 untrained machines. Sprague and Horner (1984) speculated that Student 3's poor performance on the first generalization probe following general case instruction was caused by a ritualistic pattern of inserting coins that he had developed during previous probe sessions. After receiving repeated practice on the coin insertion step during a training session between Probe Sessions 5 and 6, Student 3's performance on the generalization machines improved greatly.

Negative, or "Don't Do It," Teaching Examples

Teaching a student where and when to use a new skill or bit of knowledge does not mean that he also knows where and when *not* to use this newly learned behavior. Brian, for example, needs to learn that he should not say, "Hello, how are you?" to people that he has greeted within the past hour or so. Learners must be taught to discriminate the stimulus conditions that signal when responding is appropriate from stimulus conditions that signal when responding is inappropriate.

Instruction that includes "don't do it" teaching examples intermixed with positive examples provides learners with practice discriminating stimulus situations in which the target behavior should not be emitted (i.e., S^{Δ}s) from those conditions when the behavior is appropriate.[10] This sharpens the stimulus control necessary to master many concepts and skills (Engelmann & Carnine, 1982).

Horner and colleagues (1986) incorporated "don't do it" examples into a training program to teach four high school students with moderate to severe cognitive disabilities how to bus tables in cafeteria-style restaurants. To correctly bus a table, the student had to remove all dishes, silverware, and garbage from tabletop, chairs, and the floor under and around the table; wipe the tabletop; straighten chairs; and place dirty dishes and garbage in appropriate receptacles. In addition, the students were taught to inquire, through the use of cards, if a customer was finished with empty dishes. The three settings, one for training and two for generalization probes, differed in terms of size and furniture characteristics and configurations.

Each training trial required the student to attend to the following stimulus features of a table: (a) the presence or absence of people at a table, (b) whether people were eating at the table, (c) the amount and/or status of food on dishes, (d) the presence or absence of garbage at a table, and (e) the location of garbage and/or dirty dishes at a table. Training consisted of 30-min sessions involving six table types that represented the range of conditions most likely to be encountered in a cafeteria-style restaurant. A trainer modeled correct table busing and verbally prompted correct responding. When errors occurred, the trainer stopped the student, re-created the situation, and provided additional modeling and assistance. The six training examples consisted of four to-be-bused tables and two not-to-be-bused tables (see Table 30.1).

Generalization probe sessions were conducted in two restaurants not used for training. Each student was presented with 15 probe tables: 10 to-be-bused tables and 5 not-to-be-bused tables. The results showed a functional relation between general case instruction that included not-to-be-bused tables and "immediate and pronounced improvement in the percentage of probe tables responded to correctly" (p. 467).

Negative teaching examples are necessary when the learner must discriminate between appropriate and inappropriate conditions for a particular response. Practitioners should ask this question: Is responding *always appropriate* in the generalization setting(s)? If the answer is no, then "don't do it" teaching examples should be part of instruction.

Does the instructional setting naturally or automatically include a sufficient number and range of negative examples? The teaching situation must be analyzed to answer this important question. Practitioners may need to contrive some negative teaching examples. Practitioners should not assume the natural environment will readily reveal sufficient negative examples. Conducting training in the natural environment is no guarantee that learners will be exposed to stimulus situations that they are likely to encounter in the generalization environment after training. For example, in the study on teaching table busing, Horner and colleagues (1986) noted that, "on some days the trainer needed actively to set up one or more table types to ensure that a student had access to a table type that was not 'naturally' available" (p. 464).

"Don't do it" teaching examples should be selected and sequenced according to the degree to which they differ from positive examples (i.e., S^Ds). The most effective negative teaching examples will share many of the relevant characteristics of the positive teaching examples (Horner et al., 1988). For example, the "don't bus" tables used by Horner and colleagues (1986) shared many features with the "bus" tables (see Table 30.1). Such *minimum difference negative teaching examples* help the learner to perform the target behavior with the precision required

TABLE 30.1 Six Training Examples Used to Teach Students with Disabilities How to Bus Tables in Cafeteria-style Restaurants

Training Examples	Presence of People and Possessions	People Eating or Not Eating	Dishes: Empty/ Part/New Food	Garbage: Present or Not Present	Location of Garbage and Dishes	Correct Response
1	0 People + possessions	N/A	Partial	Present	Table, Chairs	Don't Bus
2	0 People	N/A	Partial	Present	Table, Floor, Chair	Bus
3	2 People	Eating	New food	Present	Table, Chair, Floor	Don't Bus
4	0 People	N/A	Empty	Present	Table, Floor	Bus
5	1 Person	Not eating	Empty	Present	Chair, Floor	Bus
6	2 People	Not eating	Empty	Present	Table	Bus

From "Teaching Generalized Table Bussing: The Importance of Negative Teaching Examples," by R. H. Horner, J. M. Eberhard, and M. R. Sheehan, 1986, *Behavior Modification, 10*, p. 465. Copyright 1986 by the Sage Publications, Inc. Used by permission.

by the natural environment. Minimum difference teaching examples help eliminate "generalization errors" due to overgeneralization and faulty stimulus control.

Make the Instructional Setting Similar to the Generalization Setting

> *Fresno State coach Pat Hill expects Ohio Stadium to be a new experience for the Bulldogs, who will visit for the first time. For a practice last week in FSU's stadium, Hill hired a production company to blast the Ohio State fight song—at about 90 decibels—throughout the two-hour session. "We created some noise and atmosphere to give us a feel of a live game," Hill said.*
>
> —Columbus Dispatch (August 27, 2000)[11]

A basic strategy for promoting generalization is to incorporate into the instructional setting stimuli the learner is likely to encounter in the generalization setting. The greater the similarity between the instructional environment and the generalization environment, the more likely the target behavior will be emitted in the generalization setting. The principle of stimulus generalization states that a behavior is likely to be emitted in the presence of stimuli that are very similar to the stimulus conditions in which the behavior was reinforced previously, but the behavior will likely not be emitted under stimulus conditions that differ significantly from the training stimulus.

Stimulus generalization is a relative phenomenon: The more that stimulus conditions in the generalization setting resemble the stimulus conditions present during instruction, the greater the probability that the trained response will be emitted, and vice versa. A generalization setting that differs significantly from the instructional setting may not provide sufficient stimulus control over the target behavior. Such a setting may also contain stimuli that impede the target behavior because their novelty confuses or startles the learner. Exposing the learner to stimuli during instruction that are commonly found in the generalization setting increases the likelihood that those stimuli will acquire some stimulus control over the target behavior and also prepares the learner for the presence of stimuli in the generalization setting that have the potential of hindering performance. Two tactics used by applied behavior analysts to implement this basic strategy are program common stimuli and teach loosely.

Program Common Stimuli

Program common stimuli means incorporating typical features of the generalization setting into the instructional setting. Although behavior analysts have attached a special term to this tactic, successful practitioners in many fields have long used this technique for promoting generalized behavior change. For example, coaches (like Fresno State's Coach Hill), music teachers, and theater directors hold scrimmages, mock auditions, and dress rehearsals to prepare their athletes, musicians, and actors to perform important skills in settings that include the sights, sounds, materials, people, and procedures that simulate as closely as possible those in the "real world."

Van den Pol and colleagues (1981) programmed common stimuli when they taught three young adults with disabilities how to order and eat in fast-food restaurants. The researchers used numerous items and photos from actual restaurants to make the classroom simulate the conditions found in actual restaurants. Plastic signs with pictures and names of various McDonald's sandwiches were posted on the classroom wall, a table was transformed into a mock "counter" for role-playing transactions, and the students practiced responding to 60 photographic slides taken in actual restaurants showing examples of situations customers are likely to encounter.

Bord and colleagues (2017) provide an example of programming common stimuli when teaching ice skating to a child with autism. The researchers set up a rollerblading environment to simulate the local ice-skating rink by arranging large cones in the same-sized oval as the ice-skating rink. They also included common features, such as the presence of people (e.g., center, sidelines, no people), the direction of skating (i.e., clockwise, counterclockwise) across sessions, and the presence and absence of music.

Why go to all the trouble of simulating the generalization setting? Why not just conduct instruction in the generalization setting itself to ensure that the learner experiences all of the relevant aspects of the setting? First, conducting instruction in natural settings is not always possible or practical. The resources and time required to transport students to community-based settings may not be available.

Second, community-based training may not expose students to the full range of examples they are likely to encounter later in the same setting. For example, students who receive *in situ* instruction for grocery shopping or street crossing during school hours may not experience the long lines at the checkout counters or heavy traffic patterns typical of evening hours.

Third, instruction in natural settings may be less effective and efficient than classroom instruction because the trainer cannot halt the natural flow of events to contrive an optimal number and sequence of training trials needed (e.g., Neef, Lensbower, Hockersmith, DePalma, & Gray, 1990).

Fourth, instruction in simulated settings can be safer, particularly with target behaviors that must be performed in potentially dangerous environments or that have severe consequences if performed incorrectly (e.g., Miltenberger et al., 2005), or when children or people with learning problems must perform complex procedures. If the procedures involve invading the body or errors during practice are potentially hazardous, simulation training should be used. For example, Neef, Parrish, Hannigan, Page, and Iwata (1990) had children with neurogenic bladder complications practice performing self-catheterization skills on dolls.

Programming common stimuli is a two-step process of (a) identifying salient stimuli that characterize the generalization setting(s) and (b) incorporating those stimuli into the instructional setting. A practitioner can identify possible stimuli in the generalization setting to make common by direct observation or by asking people familiar with the setting. Practitioners should conduct observations in the generalization setting(s) and note prominent features of the environment that might be important to include during training. When direct observation is not feasible, practitioners can obtain secondhand knowledge of the setting by interviewing or giving checklists to people who have

firsthand knowledge of the generalization setting—those who live, work in, or are otherwise familiar with the generalization setting(s) in question.

When possible, practitioners should use the same reinforcers during instruction as those typically produced by the target behavior in the generalization environment. Once contacted in the natural environment, the reinforcers themselves may serve as a discriminative stimuli for the target behavior (e.g., Koegel & Rincover, 1977).

If a generalization setting includes important stimuli that cannot be re-created or simulated in the instructional setting, then at least some training trials must be conducted in the generalization setting. However, as pointed out previously, practitioners should not assume that community-based instruction will guarantee students' exposure to all of the important stimuli common to the generalization setting.

Teach Loosely

Applied behavior analysts control and standardize intervention procedures to maximize their direct effects, and so the effects of their interventions can be interpreted and replicated by others. Yet restricting teaching procedures to a "precisely repetitive handful of stimuli or formats may, in fact, correspondingly restrict generalization of the lessons being learned" (Stokes & Baer, 1977, p. 358). To the extent that generalized behavior change can be viewed as the opposite of strict stimulus control and discrimination, one technique for facilitating generalization is to vary as many of the noncritical dimensions of the antecedent stimuli as possible during instruction.

Teach loosely means randomly varying noncritical aspects of the instructional setting within and across teaching sessions. This tactic has two advantages or rationales for promoting generalization. First, teaching loosely reduces the likelihood that a single or small group of noncritical stimuli will acquire exclusive control over the target behavior. A target behavior that inadvertently comes under the control of a stimulus present in the instructional setting but not always present in the generalization setting may not be emitted in the generalization setting. Here are two examples of this type of faulty stimulus control:

- ***Following teachers' directions:*** A student with a history of receiving reinforcement for complying with teachers' directions when they are given in a loud voice and accompanied by a stern facial expression may not follow directions that do not contain one or both of those noncritical variables. The discriminative stimulus for the student's compliance with teacher directions should be the content of the teacher's statements.
- ***Assembling bicycle sprocket sets:*** A new employee at the bicycle factory inadvertently learns to assemble rear sprocket sets by putting a red sprocket on top of a green sprocket and a green sprocket on top of a blue sprocket because the sprocket sets on a particular bicycle model in production on the day she was trained were colored in that fashion. However, proper assembly of a sprocket set has nothing to do with the colors of the individual sprockets; the relevant variable is the relative size of the sprockets (i.e., the biggest sprocket goes on the bottom, the next biggest on top of that one, and so on).

Systematically varying the presence and absence of noncritical stimuli during instruction greatly decreases the chance that a functionally irrelevant factor—a teacher's tone of voice or sprocket color—will acquire control of the target behavior (Kirby & Bickel, 1988).

A second rationale for loose teaching is that including a wide variety of noncritical stimuli during instruction increases the probability that the generalization setting will include some of the stimuli that were present during instruction. In this sense, loose teaching acts as a kind of catchall effort at programming common stimuli and makes it less likely that a student's performance will be impeded or "thrown off" by the presence of a "strange" stimulus.

Loose teaching applied to the previous two examples might entail the following:

- ***Following teachers' directions:*** During instruction, the teacher varies all of the factors mentioned earlier (e.g., tone of voice, facial expression), plus gives directions while standing, while sitting, from different places within the classroom, at different times of the day, while the student is alone and in groups, while looking away from the student, and so on. In each instance, reinforcement is contingent on the student's compliance with the content of the teacher's direction irrespective of the presence or absence of any of the noncritical features.
- ***Assembling bicycle sprocket sets:*** During training, the new employee assembles sprocket sets containing sprockets of widely varying colors, after receiving the component sprockets in varied sequences, when the factory floor is busy, at different times during a work shift, with and without music playing, and so forth. Irrespective of the presence, absence, or values of any of these noncritical factors, reinforcement would be contingent on correct assembly of sprockets by relative size.

Seldom used as a stand-alone tactic, loose teaching is often a recognizable component of interventions when generalization to highly variable and diverse settings or situations is desired. For example, Horner and colleagues (1986) incorporated loose teaching into their training program for table busing by systematically but randomly varying the location of the tables, the number of people at the tables, whether food was completely or partially eaten, the amount and location of garbage, and so forth. Hughes and colleagues (1995) incorporated loose teaching by varying the peer teachers and varying the locations of training sessions. Loose teaching was evident in Bord and colleagues' (2017) study when they varied the genre (e.g., rock, classical, no music) and volume (e.g., 80 db, 60 db, 40 db) of music in the instructional setting to reflect the range of music the learner might encounter in the generalization setting. Loose teaching is often a recognizable feature of language training programs that use milieu, incidental, and naturalistic teaching methods (e.g., Charlop-Christy & Carpenter, 2000; McGee, Morrier, & Daly, 1999; Warner, 1992).

Few studies evaluating the effects of using loose teaching in isolation have been reported. One exception is an experiment by Campbell and Stremel-Campbell (1982), who evaluated the effectiveness of loose teaching as a tactic for facilitating the generalization of newly acquired language by two students with moderate cognitive disabilities. The students were taught the correct use of the words *is* and *are* in "wh" questions (e.g., "What are you doing?"), yes/no reversal questions (e.g., "Is this mine?"), and statements (e.g., "These are mine?"). Each student received two 15-min language training sessions conducted within the context of other instructional activities that were part of each child's individualized education program, one during an academic task and the second during a self-help task. The student could initiate a language interaction based on the wide variety of naturally occurring stimuli, and the teacher could try to evoke a statement or question from the student by intentionally misplacing instructional materials or offering indirect prompts. Generalization probes of the students' language use during two daily 15-min free-play periods revealed substantial generalization of the language structures acquired during the loose teaching sessions.

The learner's performance of the target behavior should be established under fairly restricted, simplified, and consistent conditions, before much "looseness" is introduced. This is particularly important when teaching complex and difficult skills. Only noncritical (i.e., functionally irrelevant) stimuli should be "loosened." Practitioners should not inadvertently loosen stimuli that reliably function in the generalization setting as discriminative stimuli (S^Ds) or as "don't do it" examples (S^Δs). Stimuli known to play important roles in signaling when and when not to respond should be systematically incorporated into instructional programs as teaching examples. A stimulus condition that may be functionally irrelevant for one skill may be a critical S^D for another skill.

Taking the notion of varying noncritical aspects of the instructional setting and procedures to its logical limit, Baer (1999) offered the following advice for loose teaching:

- Use two or more teachers.
- Teach in two or more places.
- Teach from a variety of positions.
- Vary your tone of voice.
- Vary your choice of words.
- Show the stimuli from a variety of angles, using sometimes one hand and sometimes the other.
- Have other persons present sometimes and not other times.
- Dress quite differently on different days.
- Vary the reinforcers.
- Teach sometimes in bright light, sometimes in dim light.
- Teach sometimes in noisy settings, sometimes in quiet ones.
- In any setting, vary the decorations, vary the furniture, and vary their locations.
- Vary the times of day when you and everyone else teach.
- Vary the temperature in the teaching settings.
- Vary the smells in the teaching settings.
- Within the limits possible vary the content of what's being taught.
- Do all of this as often and as unpredictably as possible. (p. 24)

Of course, Baer (1999) was not suggesting that a teacher needs to vary all of these factors for every behavior taught. But building a reasonable degree of "looseness" into teaching is an important element of a teacher's overall effort to program for generalization rather than train and hope.

Comparing Teach Enough Examples, Program Common Stimuli, and Teach Loosely. Teach enough examples, program common stimuli, and teach loosely are three related tactics for promoting generalized outcomes based upon the principle of stimulus control. Each strategy entails the systematic selection, presentation, and variation of antecedent stimuli during instruction. The definition and rationale for each tactic and examples of stimuli that might be included in a program to teach a novice driver how to parallel park are shown in Table 30.2.

The three tactics differ from one another on the basis of how critical a particular antecedent stimulus is to the target behavior and the specificity with which the stimulus is identified and controlled during instruction. The tactics can be compared and contrasted as illustrated in Figure 30.7. At one end of the importance and specificity continuum is *teach loosely*, a tactic in which noncritical stimuli are randomly varied during and across teaching sessions. At the other end lies *teach enough examples*, which requires the identification and precise presentation of instructional examples representing critical features found in the generalization setting/situation. Relatively speaking, *programming common stimuli* falls in the middle of the continuum and overlaps with the tactics on either side.

Judging where a given stimulus belongs on the importance continuum can be puzzling. When teaching parallel parking, should the amount and flow of traffic be incorporated into specific teaching examples or programmed as a common stimulus? Should passengers talking to the driver or one another be programmed as a common stimulus or varied among a collection of other noncritical stimuli as part of loose teaching? A good case could be made for either answer in both examples. That there are no right and wrong answers to these questions poses no hindrance to planning and delivering instruction for generalized behavior change. What is important is taking all potentially relevant stimuli into account. We recommend practitioners integrate stimuli that fall in the overlaps between these strategies with this conservative approach: Build stimuli that might be teaching examples or common stimuli into the instructional plan as specific teaching examples; program as common stimuli those that fall between program common stimuli and teach loosely.

Maximize Contact with Reinforcement in the Generalization Setting

Even though the learner emits a newly acquired target behavior in a generalization setting, generalization and maintenance will be short-lived if the behavior does not contact naturally existing contingencies of reinforcement. Five of the 13 tactics for promoting generalized behavior change described in this

TABLE 30.2 Teach Enough Examples, Program Common Stimuli, and Teach Loosely Compared

Strategy	Rationale	Examples: Teaching Parallel Parking
Teach enough examples—Select and teach examples representing the full range of critically relevant stimulus variations and response requirements typically found in the generalization setting.	To increase the likelihood of correct responses to untaught examples encountered in the generalization setting.	• Available parking spaces of different lengths • Parked car at front, back, or both ends of space • Driving own car, partner's car, rental car • In a hurry/late for appointment • "No Parking" sign or hood over parking meter • Passenger questioning or challenging driver's ability to parallel park
Program common stimuli—Include stimuli in the instructional setting that are known or are likely to exist in the generalization setting.	To increase the likelihood of setting/situation generalization.	• Heavy traffic, little or no traffic • One or more passengers in car • Pedestrians on sidewalk • Radio on/off • Daytime or at night
Teach loosely—Vary functionally irrelevant stimuli within and across teaching sessions.	To (a) minimize the likelihood of an irrelevant stimulus acquiring stimulus control over the target behavior and (b) program common stimuli to the extent the learner encounters any of the "loosened" stimuli in the generalization settings.	• Music, sports, or news on radio at different volumes • Passenger(s) talking to driver or one another • Weather: sun, rain, fog, snow • Expensive car(s) parked in front and/or back of space • Police officer or police car nearby

chapter involve some form of arranging or contriving for the target behavior to be reinforced in the generalization setting.

Teach Behavior to Levels Required by Natural Contingencies

Baer (1999) noted a common mistake by practitioners: failing to teach the behavior change to performance levels sufficient to meet natural contingencies of reinforcement.

> Behavior changes that seem to need generalization may only need better teaching. Try making the students fluent, and see if they still need further support for generalization. Fluency may consist of any or all of the following: high rate of performance, high accuracy of performance, fast latency given the opportunity to perform, and strong response. (p. 17)

A new behavior may occur in the generalization setting but fail to make contact with naturally existing contingencies of reinforcement. Ensuring that a new behavior change meets naturally existing contingencies of reinforcement may require enhancing the learner's performance of the target behavior in one or more of these dimensions: accuracy, rate, duration, latency, magnitude, and topography. For example, when given a worksheet to complete at his desk, a student's behavior that is consistent with the following dimensions is unlikely to contact reinforcement for completing the task, even if the student has the ability to complete each worksheet item accurately.

- ***Latency too long.*** A student who spends 5 minutes "daydreaming" before he begins reading the directions may not finish in time to obtain reinforcement.
- ***Rate too low.*** A student who needs 5 minutes to read the directions for an independent seatwork assignment that his peers read in less than 1 minute may not finish in time to obtain reinforcement.
- ***Duration too brief.*** A student who can work without direct supervision for only 5 minutes at a time will not be able to complete any task requiring more than 5 minutes of independent work.

The solution for this kind of generalization problem, while not always simple, is straightforward. Instruction must bring target behavior performance to levels commensurate with the

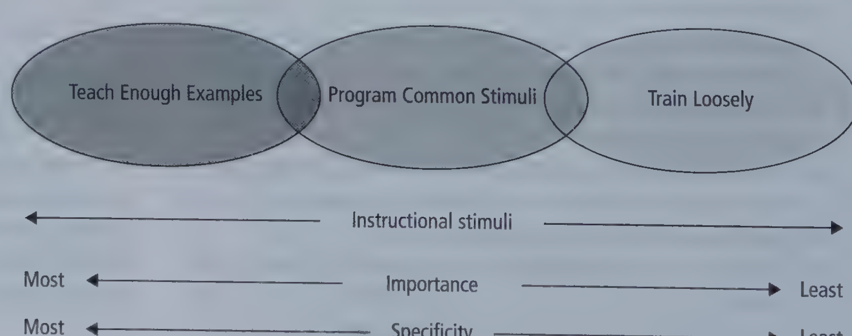


Figure 30.7 Relative importance and specificity of instructional stimuli selected for the generalization-promoting tactics: teach enough examples, program common stimuli, and teach loosely.

naturally occurring contingencies in the generalization setting. A good generalization plan specifies the performance necessary to meet naturally existing criteria for reinforcement.

Program Indiscriminable Contingencies

The clear, predictable, and immediate consequences that are so effective in helping learners acquire new behavior can work against generalized responding. This is most likely to occur when a newly acquired skill has not yet contacted naturally existing contingencies of reinforcement, and the learner can discriminate when the instructional contingencies are absent in the generalization settings. When the presence or absence of the controlling contingencies can be detected in the generalization setting ("The game's off, no need to respond here/now."), the learner may stop responding in the generalization setting, and the behavior change the practitioner worked so hard to develop may cease to occur before it can contact the naturally existing contingencies of reinforcement.

An **indiscriminable contingency** is one in which the learner cannot discriminate whether or not the next response will produce reinforcement. As a tactic for promoting generalization and maintenance, programming indiscriminable contingencies involves contriving a contingency in the generalization setting in which (a) reinforcement is contingent on some, but not all, occurrences of the target behavior, and (b) the learner is unable to predict which responses will produce reinforcement.

The basic rationale for programming indiscriminable contingencies is to keep the learner responding often enough and long enough in the generalization setting for the target behavior to make sufficient contact with the naturally existing contingencies of reinforcement. From that point on, the need to program contrived contingencies to promote generalization will be moot. Applied behavior analysts use two related techniques to program indiscriminable contingencies: intermittent schedules of reinforcement and delayed rewards.

Intermittent Schedules of Reinforcement. A newly learned behavior often must occur repeatedly over a period of time in the generalization setting before it contacts a naturally existing contingency of reinforcement. During that time, an extinction condition exists for responses emitted in the generalization setting. The current or most recent schedule of reinforcement for a behavior in the instructional setting plays a significant role in how many responses will be emitted in the generalization setting prior to reinforcement. Behaviors that have been under continuous schedules of reinforcement (CRF) show limited response maintenance under extinction. When reinforcement is no longer available, responding is likely to decrease rapidly to pre-reinforcement levels. In contrast, behaviors with a history of intermittent schedules of reinforcement often continue to be emitted for relatively long periods after reinforcement is no longer available (e.g., Dunlap & Johnson, 1985; Hoch, McComas, Thompson, & Paone, 2002).

Koegel and Rincover (1977, Experiment II) demonstrated the effects of intermittent schedules of reinforcement on response maintenance in a generalization setting. The participants were six boys diagnosed with autism and severe to profound cognitive disabilities, ages 7 to 12 years, who had participated in a previous study and had shown generalized responding in the extra-therapy setting used in that experiment (Rincover & Koegel, 1975). As in Experiment I by Koegel and Rincover (1977) described earlier in this chapter, one-on-one training trials were conducted with each child and the trainer seated at a table in a small room, and generalization trials were conducted by an unfamiliar adult standing outside on the lawn, surrounded by trees. The target behavior was (a) nonverbal imitation (e.g., raising arm) in response to an imitative model and the verbal instruction "Do this" and (b) touching a body part in response to verbal instructions such as "Touch your nose." After acquiring an imitation response, each child was given additional trials on one of three randomly chosen schedules of reinforcement: CRF, fixed ratio (FR) 2, or FR 5. After these additional training trials, each child was taken outside to assess response maintenance. Once outside, trials were conducted until the child's correct responding had decreased to 0% or maintained at 80% correct or above for 100 consecutive trials.

Behaviors that were most recently on a CRF schedule in the instructional setting underwent extinction quickly in the generalization setting (see Figure 30.8). Generalized responding occurred longer for the FR 2 trained behavior, and longer still for behavior that had been shifted to an FR 5 schedule in the instructional setting. The results showed clearly that the schedule of reinforcement in the instructional setting had a predictable effect on responding in the absence of reinforcement in the generalization setting: The thinner the schedule in the instructional setting, the longer the response maintenance in the generalization setting.

The defining feature of all intermittent schedules of reinforcement is that only some responses are reinforced, which means, of course, that some responses go unreinforced. One possible explanation for the maintenance of responding during periods of extinction for behaviors developed under intermittent schedules is the relative difficulty of discriminating that reinforcement is no longer available. Thus, the unpredictability of an intermittent schedule may account for the maintenance of behavior after the schedule is terminated.

Although *all indiscriminable contingencies of reinforcement involve intermittent schedules, not all schedules of intermittent reinforcement are indiscriminable.* For example, although the FR 2 and FR 5 schedules of reinforcement used by Koegel and Rincover (1977) were intermittent, many learners would soon be able to discriminate whether reinforcement would follow their next response. In contrast, a student whose behavior is being supported by a variable ratio (VR) 5 schedule of reinforcement cannot determine if his next response will be reinforced.

Delayed Rewards. Stokes and Baer (1977) suggested that not being able to discriminate the time when a behavior will be reinforced is similar to not being able to discriminate whether the next response will be reinforced. They cited an experiment by Schwarz and Hawkins (1970) in which each day after school a sixth-grade girl was shown videotapes of her behavior in that day's math class and received praise and token reinforcement for improvements in her posture, reducing the number of times she touched her face, and speaking with sufficient volume to be heard by others. Reinforcement after school was contingent

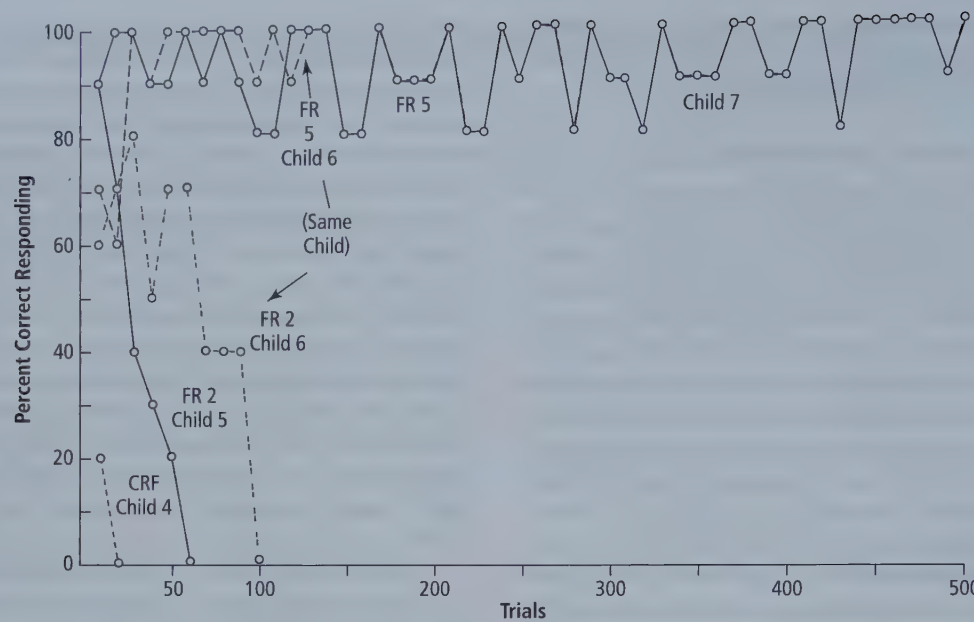


Figure 30.8 Percent of correct responses by three children in a generalization setting as a function of the schedule of reinforcement used during the final sessions in an instructional setting.

From "Research on the Differences between Generalization and Maintenance in Extra-Therapy Responding" by R. L. Koegel and A. Rincover, 1977, *Journal of Applied Behavior Analysis, 10,* p. 8. Reproduced with permission of John Wiley & Sons, Inc.

on behaviors emitted during math class only, but comparable improvements were noted in spelling class as well. The generalization data were taken from videotapes of the girl's behavior in spelling class and were never shown to her. Stokes and Baer hypothesized that because reinforcement was delayed (the behaviors that produced praise and tokens were emitted during math class but were not rewarded until after school), it may have been difficult for the student to discriminate when improved performance was required for reinforcement. They suggested that the generalization across settings of the target behaviors may have been a result of the indiscriminable nature of the response-to-reinforcement delay.

Delayed rewards and intermittent schedules of reinforcement are alike in two ways: (a) Reinforcement is not delivered each time the target behavior is emitted (only some responses are followed by reinforcement), and (b) there are no clear stimuli to signal which responses will produce reinforcement. Delayed reward contingencies differ from intermittent reinforcement in that instead of the consequence immediately following an occurrence of the target behavior, the reward is provided after a period of time has elapsed (i.e., a response-to-reward delay). Receiving the delayed reward is contingent on the learner having performed the target behavior in the generalization setting *during an earlier time period.* With an effective delayed reward contingency, the learner cannot discriminate when (and perhaps where, depending on the details of the contingency) the target behavior must be emitted to obtain a reward later. As a result, the learner's best chance to receive the reward later is to "be good all day" (Fowler & Baer, 1981).

Two similar studies by Freeland and Noell (1999, 2002) investigated the effects of delayed rewards on the maintenance of students' mathematics performance. Participants in the second study were two third-grade girls who had been referred by their teacher for help with mathematics. That target behavior for both students was writing answers to single-digit addition problems with sums to 18. The researchers used a multiple-treatment reversal design to compare the effects of five conditions on the number of correct digits written as answers to single-digit addition problems during daily 5-min work periods (e.g., writing "11" as the answer to "5 + 6 = ?" counted as two digits correct).

- *Baseline:* Green worksheets; no programmed consequences. Students were told they could attempt as many or as few problems as they wanted.
- *Reinforcement:* Blue worksheets with a goal number at the top indicating the number of correct digits needed to choose a reward from the "goody box." Each student's goal number was the median number of correct digits on the last three worksheets. All worksheets were graded after each session, and reinforcement was contingent on meeting the performance goal.
- *Delay 2:* White worksheets with goal number; after every two sessions, one of the two worksheets completed by each student was randomly selected for grading; reinforcement was contingent on meeting highest median of three consecutive sessions up to that point in the study.
- *Delay 4:* White worksheets with goal number and same procedures as Delay 2 except that worksheets were not graded until four sessions had been completed, at which time one of each student's previous four worksheets was randomly selected and graded.

- *Maintenance:* White worksheets with goal number as before; no worksheets were graded and no feedback or rewards for performance were given.

The different-colored worksheets used for each condition in this study made it easy for the students to predict the likelihood of reinforcement. A green worksheet meant no feedback or other consequences no matter how many correct digits were written. A blue worksheet indicated that meeting the performance goal would earn a reward from the "goody box." However, meeting the performance criterion on a white worksheet sometimes produced reinforcement. This study demonstrates the importance of having contingencies in the instructional setting "look like" the contingencies in effect in the generalization setting(s) in two ways: (a) Both students showed large decreases in performance when baseline conditions were reinstated, and stopped responding altogether during a second return to baseline; and (b) the students continued completing math problems at a high rate during the maintenance condition, even though no reinforcement was provided (see Figure 30.9).

When the delayed (indiscriminable) contingencies were implemented, both students demonstrated levels of correct responding at or above levels during the reinforcement phase. When the students were exposed to maintenance conditions, Amy maintained high levels of responding for 18 sessions with variable performance over the final six sessions, and Kristen showed a gradually increasing rate of performance over 24 sessions. The results demonstrated that behavior with an indiscriminable contingency can be maintained at the same rate as with a predictable schedule, and with greater resistance to extinction.

Delayed consequences have been used to promote the setting/situation generalization and response maintenance of a wide range of target behaviors, including academic and vocational tasks by individuals with autism (Dunlap, Koegel, Johnson, & O'Neill, 1987); young children's toy play, social initiations, and selection of healthy snacks (R. A. Baer, Blount, Dietrich, & Stokes, 1987; R. A. Baer, Williams, Osnes, & Stokes, 1984; Osnes, Guevremont, & Stokes, 1986); restaurant trainees' responding appropriately to coworkers' initiations (Grossi et al., 1994); and performance on academic tasks (Brame, 2001; Theodore, Bray, Kehle, & Jenson, 2001).

Effective use of delayed consequences can reduce (or even eliminate in some instances) the learner's ability to discriminate when a contingency is or is not in effect. As a result, the learner needs to "be good" (i.e., emit the target behavior) all the time. If an effective contingency can be made indiscriminable across

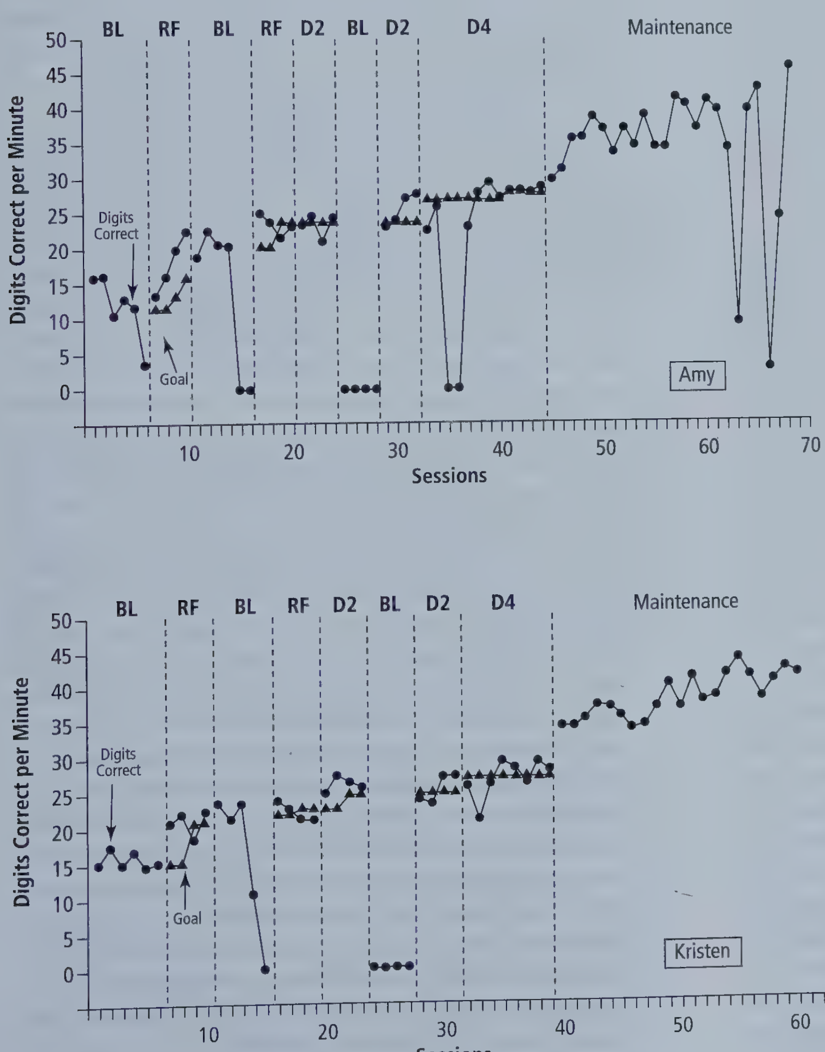


Figure 30.9 Number of correct digits per minute by two third-grade students while answering math problems during baseline (BL), reinforcement contingent on meeting the performance goal on one randomly selected worksheet after each session (RF), reinforcement contingent on meeting the performance goal on one randomly selected worksheet after every two (D2) or four (D4) sessions, and maintenance conditions.

From "Programming for Maintenance: An Investigation of Delayed Intermittent Reinforcement and Common Stimuli to Create Indiscriminable Contingencies," by J. T. Freeland and G. H. Noell, 2002, *Journal of Behavioral Education, 11*, p. 13. Copyright 2002 by Human Sciences Press. Reprinted by permission.

settings and target behaviors, the learner will also have to "be good" everywhere, with all of his or her relevant skills.

Here are two examples of how the supervisor of an office cleaning crew could use an indiscriminable contingency to reinforce quality work by an employee.

- At shift's end, the employee pulls from the supervisor's hat a slip of paper on which a room number and item to be cleaned are written (e.g., #566, carpet vacuumed). If the randomly selected item has been cleaned properly, the employee earns a reward.
- The employee initials and leaves a sticky note on the door of each room she has cleaned. The supervisor collects the notes and initials those from properly cleaned rooms. At shift's end, the employee draws a sticky-note from all the notes she left during that shift. If the randomly drawn note contains the supervisor's initials, she gets a reward.

Following are five examples of classroom applications of indiscriminable contingencies involving delayed rewards. Each application features an interdependent group contingency that makes group rewards contingent on the performance of randomly selected students (see Chapter 28).

- *Spinners and Dice I.* Teachers can make academic seatwork periods more effective with spinners and dice. Every few minutes (e.g., on a variable interval [VI] 5-min schedule), the teacher (a) randomly selects a student's name, (b) walks to that student's desk and has the student twirl a spinner or roll a die, (c) counts backward from the worksheet problem or item the student is currently working on by the number shown on the spinner or die, and (d) gives a token to the student if that problem or item is correct. Students who immediately begin to work on the assignment and work quickly but carefully throughout the seatwork period are most likely to obtain reinforcers under this indiscriminable contingency.
- *Spinners and Dice II.* In a study by L. D. Heward, Park, and Alber-Morgan (2008), second-graders completed a variety of academic tasks on the board. At the end of the timed work session, one student spun a spinner that determined the subject area (reading, time telling, social studies, or language). A second student then rolled dice that indicated the item number. Each student who correctly answered the randomly selected item earned a reward.
- *Story Fact Recall Game.* Many elementary teachers devote 20 to 30 minutes per day to sustained silent reading (SSR), a period when students can read silently from books of their choice (Temple, Ogle, Crawford, & Freppon, 2018). An indiscriminable contingency can encourage students to read with purpose during SSR. At the end of the SSR period, the teacher asks several randomly selected students a question about the book they are reading. A student reading Chapter 3 of Elizabeth Winthrop's *The Castle in the Attic* might be asked, "What did William give the Silver Knight to eat?" (Answer: bacon and toast). Correct answers are praised by the teacher, are applauded by the class, and earn a marble in a jar toward a reward for the whole class. Students do not know when they will be called on or what they might be asked (Brame, Heward, & Bicard, 2002).
- *The Jars.* The teacher arranges three jars containing slips of paper or cards indicating different behaviors, performance criteria, and rewards. At the conclusion of a preannounced time, the teacher randomly selects from one jar the subject area (e.g., math, science, social studies) or classroom rule (e.g., participation, no talk-outs), from a second jar the performance criteria (e.g., 70%, 80%, or 90% accuracy; 10, 15, or 20 problems completed), and from a third jar the reward to be earned if the performance criterion is met. A fourth jar might include names of individual students or subgroups within the class who must meet the performance criterion for the class to earn the reward. Several studies have found "the jars" to be a fun and effective method to improve and generalize students' academic performance and compliance with classroom rules (e.g., Kelshaw-Levering, Sterling-Turner, Henry, & Skinner, 2000; Theodore et al., 2001; Zibreg Hargis, Pattie, Maheady, Budin, & Rafferty, 2017).
- *Numbered Heads Together.* Having small groups of students work together on a joint learning activity can be effective, but teachers should use procedures that motivate all students to participate. A collaborative learning technique called Numbered Heads Together ensures that all students actively participate (Hunter et al., 2015; Kagan & Kagan, 2009). Students are seated in heterogeneous groups of three or four, and each student is given the number 1, 2, 3, or 4. The teacher asks the class a question, and each group discusses the problem and comes up with an answer. Next, the teacher randomly selects a number from 1 to 4 and then calls on one or more students with that number to answer. It is important that every person in the group knows the answer to the question. This strategy promotes cooperation within the group rather than competition. Because all students must know the answer, group members help each other understand not only the answer but also the how and why behind it. Finally, this strategy encourages individual responsibility.
- *Intermittent Grading.* Most students do not receive sufficient practice writing, and when students do write, the feedback they receive is often ineffective. One reason for this may be that providing detailed feedback on a daily composition by each student in the class requires more time and effort than even the most dedicated teacher can give. A procedure called *intermittent grading* offers one solution to this problem (Heward, Heron, Gardner, & Prayzer, 1991). Students write for 10 to 15 minutes each day, but instead of reading and evaluating every student's paper, the teacher provides detailed feedback on a randomly selected 20% to 25% of students' daily compositions. The students whose papers were graded earn points based on individualized performance criteria, and bonus points are given to the class contingent on the quality of the selected and graded papers (e.g., if the authors of four of the five papers that were graded met their individual criteria). Students' papers that were graded can be used as a source of instructional examples for the next lesson.

The effectiveness of a delayed reward tactic in promoting generalization and maintenance rests on (a) the indiscriminability of the contingency (i.e., the learner cannot tell exactly when emitting the target behavior in the generalization setting will produce a reward at a later time) and (b) the learner understanding the relation between emitting the target behavior and receiving a reward at a later time for doing so. Box 30.2 describes how indiscriminable contingencies can be used to foster sustainable practices such as recycling and energy conservation.

Guidelines for Programming Indiscriminable Contingencies. Practitioners should consider these guidelines when implementing indiscriminable contingencies:

- Use continuous reinforcement during the initial stages of developing new behaviors or when strengthening little-used behaviors.
- Systematically thin the schedule of reinforcement based on the learner's performance (see Chapter 13). Remember that the thinner the schedule of reinforcement is, the more indiscriminable it is (e.g., an FR 8 schedule is more indiscriminable than an FR 2 schedule); and variable schedules of reinforcement (e.g., VR and VI schedules) are more indiscriminable than fixed schedules are (e.g., FR and fixed interval [FI] schedules).
- When using delayed rewards, begin by delivering the reinforcer immediately following the target behavior and gradually increase the response-to-reinforcement delay.
- Each time a delayed reward is delivered, explain to the learner that he is receiving the reward for specific behaviors he performed earlier. This helps to build and strengthen the learner's understanding of the rule describing the contingency. A delayed rewards intervention may not be effective with some learners who have significant intellectual disabilities.

Set Behavior Traps

Some contingencies of reinforcement are especially powerful, producing substantial and long-lasting behavior changes. Baer and Wolf (1970b) called such contingencies **behavior traps**. Using a mouse trap as an analogy, they described how a householder has only to exert a relatively small amount of behavioral control over the mouse—getting the mouse to smell the cheese—to produce a behavior change with considerable (in this case, complete) generalization and maintenance.

> A householder without a trap can, of course, still kill a mouse. He can wait patiently outside the mouse's hole, grab the mouse faster than the mouse can avoid him, and then apply various forms of force to the unfortunate animal to accomplish the behavioral change desired. But this performance requires a great deal of competence: vast patience, super-coordination, extreme manual dexterity, and a well-suppressed squeamishness. By contrast, a householder with a trap needs very few accomplishments: If he can merely apply the cheese and then leave the trap where the mouse is likely to smell that cheese, in effect he has guaranteed general(ized) change in the mouse's future behavior.
>
> The essence of a trap, in behavioral terms, is that *only a relatively simple response is necessary to enter the trap, yet once entered, the trap cannot be resisted in creating general behavior change*. For the mouse, the entry response is merely to smell the cheese. Everything proceeds from there almost automatically. (p. 321, emphasis added)

Behavioral trapping is a fairly common phenomenon that everyone experiences from time to time. Behavior traps are particularly evident in the activities students just cannot get (or do) enough of. The most effective behavior traps share four essential features: (a) They are "baited" with virtually irresistible reinforcers that "lure" the student to the trap; (b) only a low-effort response already in the student's repertoire is necessary to enter the trap; (c) interrelated contingencies of reinforcement inside the trap motivate the student to acquire, extend, and maintain targeted academic and/or social skills; and (d) they can remain effective for a long time because students show few, if any, satiation effects.

Consider the case of the "reluctant bowler." A young man is persuaded to fill in as a substitute for a friend's bowling team. He has always regarded bowling as uncool. And bowling looks so easy on television that he does not see how it can be considered a real sport. Nevertheless, he agrees to go, just to help out this one time. During the evening he learns that bowling is not as easy as he had always assumed (he has a history of reinforcement for athletic challenges) and that several people he would like to get to know are avid bowlers (i.e., it is a mixed doubles league). Within a week he has purchased a custom-fitted bowling ball, a bag, and shoes; practiced twice on his own; and signed up for the next league.

The reluctant bowler example illustrates the fundamental nature of behavior traps: easy to enter and difficult to exit. Some naturally existing behavior traps can lead to maladaptive behaviors, such as alcoholism, drug addiction, and juvenile delinquency. The everyday term *vicious circle* refers to the natural contingencies of reinforcement that operate in destructive behavior traps. Practitioners, however, can learn to create behavior traps that help students develop positive, constructive knowledge and skills. Alber and Heward (1996) provided guidelines for creating successful traps and gave the following example of an elementary teacher creating a behavior trap that took advantage of a student's penchant for playing with baseball cards.

> Like many fifth graders struggling with reading and math, Carlos experiences school as tedious and unrewarding. With few friends of his own, Carlos finds that even recess offers little reprieve. But he does find solace in his baseball cards, often studying, sorting and playing with them in class. His teacher, Ms. Greene, long ago lost count of the number of times she had to stop an instructional activity to separate Carlos and his beloved baseball cards. Then one day, when she approached Carlos' desk to confiscate his cards in the middle of a lesson on alphabetization, Ms. Greene discovered that Carlos had already alphabetized all the left-handed pitchers in the National League! Ms. Greene realized she'd found the secret to sparking Carlos' academic development.

BOX 30.2

Fostering Earth-friendly Behavior with Clueless Contingencies

Earth's average temperature in 2016 was the warmest ever recorded, and when 2019 began, the ten warmest years on record had all occurred since 1998 (National Aeronautics and Space Administration [NASA], 2019; National Oceanic and Atmospheric Administration [NOAA], 2019). This global warming has contributed to melting ice sheets and rising sea levels, an increased frequency of catastrophic floods, droughts, and forest fires, and has threatened the survival of some species (Intergovernmental Panel on Climate Change, 2018). Most of the warming results from increased concentrations of carbon dioxide (CO_2) and other greenhouse gases in the atmosphere due to human activity since the mid-20th century, primarily the burning of fossil fuels (Intergovernmental Panel on Climate Change, 2014).

Although some climate change deniers persist in complete disregard of scientific facts, the challenge has shifted from convincing people climate change is real to getting them to change their behavior in response to it (Thompson, 2010). Even if all fossil fuel emissions ceased today, our climate would continue warming for decades (Marcott, Shakun, Clark, & Mix, 2013). Successfully adapting to that reality will require massive shifts to renewable energy sources and changes in transportation infrastructure, food production, environmental conservation, and economic practices on a global scale far beyond the behavior changes any of us can make now. But the changes we do make now can provide a *behavioral wedge* that gives society time to discover the technological fixes and implement the policy changes needed to make realities of carbon neutrality and a truly sustainable society (Dietz, Gardner, Gilligan, Stern, & Vandenbergh, 2009).

Rewarding Green Behavior

Richard Thaler (2012), economist and co-author of *Nudge: Improving Decisions About Health, Wealth, and Happiness*, began a *New York Times* essay, "Making Good Citizenship Fun," as follows: "Governments typically use two tools to encourage citizens to engage in civic behavior like paying their taxes, driving safely, or recycling their garbage: exhortation and fines. These efforts are often ineffective. So it might be a good time to include positive reinforcement." Thaler then described several examples of reinforcement-based government campaigns to promote behavior that many people avoid, including obeying the speed limit in Sweden, paying taxes in China, and picking up dog feces off the sidewalks in Taiwan.

Behavior analysts should be delighted that efforts to reward good behavior are catching on as a way to ameliorate social problems. Lotteries and related incentive systems are practically effective and conceptually consistent with behavioral principles. The overall impact of most such programs, however, is limited by a restricted footprint (limited setting, reward a single type of green behavior, reward individuals,) and predictability: "The game's off. No need to respond now."

An indiscriminable contingency approach to enhancing existing green initiatives can be designed, either for individuals or groups, to address multiple responses in many settings with many rewards. A green behavior change program featuring a well-designed indiscriminable contingency presents a radically different and enticing "if-then" contingency: When you can't tell which of several green behaviors performed exactly where or when will produce a reward, the best strategy for optimizing reward is acting green in every way, everywhere, all the time.

Let's Play Conservation Clue

Conservation Clue is a whodunit competition like the classic board game Clue™, but instead of trying to evade getting caught, players leave clues of their good green deeds all over the place. Examples of target behaviors, measures, rewards, and gaming fun that could be part of Conservation Clue games played in residential, work, and community settings are shown in Figure A.

1. **Target Green Behaviors.** Select oft-repeated behaviors such as turning off lights, unplugging electronics, or taking public transit. Once-only behaviors such as buying a plug-in hybrid or installing solar panels do not lend themselves to this approach (though they can be incentivized in other ways; see Chance & Heward, 2010).
2. **Determine How to Detect and Measure the Behaviors.** Some behaviors leave products that are readily measured, such as the number of outlets or bulbs in use, or the weight or volume in a recycle bin. Other behaviors must be recorded as they occur, such as punches on a bus pass or tickets dispensed for bringing a reusable bag. Recruiting people to monitor and measure green actions can pay off with additional prompts, modeling, or incidental social reinforcement for the targeted behaviors. Members of student or community groups with environmental and sustainability missions would gladly serve as "Conservation Clue Detectives."
3. **Ready the Playing Field.** Just as athletes play better on a well-groomed diamond or rink, Conservation Clue gamers will behave greener in well-prepared environments. Preparation entails two primary actions: minimizing barriers to responding (e.g., putting recycling bins where they are easy to access) and alerting participants to the opportunity to play (in other words, providing prompts to engage in green behavior).
4. **Select Rewards.** Examples of simple, relatively inexpensive rewards for a campus-based game of Conservation Clue include school insignia wear, bookstore discounts, lunch with the dean, concert tickets, and gift certificates for local green businesses. Virtual objects such as stars, trophies, and green leaves that gamers receive and collect online can be effective and fun rewards (Pritchard, 2010; Twyman, 2010).

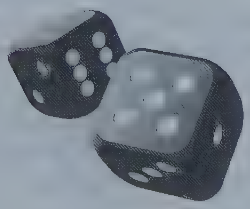
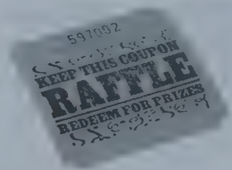
Settings	 **Where We Live**	 **Where We Work**	 **Where We Shop**
Behaviors	• unplug • reduce water use • reduce heat and A/C • reduce solid waste (recycle & compost) • reuse containers	• recycle paper/waste • carpool/take public transportation • power down idle equipment • use vehicles efficiently • reuse materials	• buy locally sourced food and goods • buy products in bulk or reduced packaging • buy items made with recycled content • shop with reusable bags
Measures	• plugs in use • bulb counts • utility bills • meter readings • weight of recyclables • purchase receipts	• public transit receipts • GPS data from fleet vehicles • employees initial reused materials • green officer verification	• store receipts • bar code scans • cashier verification
Rewards	• discounts at local shops • recycle bin stickers • utility bill discounts • point accrual • virtual rewards	• recognition • extra breaks • lunch with the boss • company gear • cash bonuses	• reduced cost items • gift cards • charity donations in shopper's name • electronic fanfare • public posting
The Game	• jar drawings • dice rolls • dart throws 	• spinners • wheels of chance • lotteries 	• random item on sales receipts half-off • raffles 

Figure A Target behaviors, measures, rewards, and game methods that could be used in Conservation Clue, played where we live, work, and shop.

5. **Make It a Game.** The centerpiece of Conservation Clue is a game-like procedure for randomly determining which green behaviors, emitted where and when, will earn what rewards. The jars technique described earlier is one of countless ways of implementing this multidimensional indiscriminable contingency. Game rules and materials—spinners or dice or dart boards—should be selected to suit the setting and players' interests; the crucial element is that multiple forms of green behavior are rewarded in a way that is utterly unpredictable . . . and fun.

(continued)

BOX 30.2 (continued)

6. **Evaluate, Revise, and Play Again.** In the spirit of the science from which it is derived, Conservation Clue should be an empirical endeavor. Gather baseline data on target behaviors against which to compare any gains that subsequently come as a result of the playing. The game includes many variables to monitor and experiment with, such as which green behaviors are more malleable, which rewards are more preferred, and what time intervals between drawings are optimal. Player satisfaction should be assessed and suggestions for improvements sought.

Conservation Clue is one way to back good intentions with consequential muscle. The behavior changes may seem trivial at first. But many small actions can redefine what it means to act locally and have global impact. Imagine Conservation Clue being played by the residents of London or Shanghai, drivers on the New Jersey Turnpike, or employees of a major corporation with a global footprint. What if Walmart's more than 2 million employees at the company's 4253 stores around the globe engaged in a full-fledged game of Conservation Clue? What if they then invited the 200 million customers who shop at their stores each week to join in the fun?

Adapted from "Sustaining Sustainability with Clueless Contingencies," by W. L. Heward and J. W. Kimball, 2013, *Sustain Magazine*, *28*, 4–15. Used with permission.

Carlos was both astonished and thrilled to learn that Ms. Greene not only let him keep his baseball cards at his desk, but also encouraged him to "play with them" during class. Before long, Ms. Greene had incorporated baseball cards into learning activities across the curriculum. In math, Carlos calculated batting averages; in geography, he located the hometown of every major leaguer born in his state; and in language arts, he wrote letters to his favorite players requesting an autographed photo. Carlos began to make significant gains academically and an improvement in his attitude about school was also apparent.

But school became really fun for Carlos when some of his classmates began to take an interest in his knowledge of baseball cards and all the wonderful things you could do with them. Ms. Greene helped Carlos form a classroom Baseball Card Club, giving him and his new friends opportunities to develop and practice new social skills as they responded to their teacher's challenge to think of new ways to integrate the cards into the curriculum. (p. 285)

Ask People in the Generalization Setting to Reinforce the Target Behavior

The problem may be simply that the natural community [of reinforcement] is asleep and needs to be awakened and turned on.

—Donald M. Baer (1999, p. 16)

Sometimes a potentially effective contingency of reinforcement in the generalization setting is not operating in a form available to the learner no matter how often or well she performs the target behavior. The contingency is there, but dormant. One solution for this kind of problem is to inform key people in the generalization setting of the value and importance of their attention to the learner's efforts to acquire and use new skills and ask them to help.

For example, a special education teacher who has been helping a student learn how to participate in class discussions by providing repeated opportunities for practice and feedback in the resource room could inform the teachers in the general education classes that the student attends about the behavior change program and ask them to look for and reinforce any reasonable effort by the student to participate in their classrooms. A small amount of contingent attention and praise from these teachers may be all that is needed to achieve the desired generalization of the new skill.

This simple but often effective technique for promoting generalized behavior was evident in the study by Stokes and colleagues (1974) in which staff members responded to the children's waving response by saying, "Hello, (name)." Approximately 20 such generalization probes were conducted each day with each child.

Williams, Donley, and Keller (2000) gave mothers of two preschool children with autism explicit instructions on providing models, response prompts, and reinforcement to their children who were learning to ask questions about hidden objects (e.g., "What's that?" "Can I see it?").

Contingent praise and attention from significant others can add to the effectiveness of other strategies already in place in the generalization environment. Broden, Hall, and Mitts (1971) showed this in the self-monitoring study described in Chapter 29. After self-recording had improved eighth-grader Liza's study behavior during history class, the researchers asked Liza's teacher to praise Liza for study behavior whenever he could during history class. Liza's level of study behavior during this Self-Recording Plus Praise condition increased to a mean of 88% and was maintained at levels nearly as high during a subsequent Praise Only condition (see Figure 29.3).

Teach the Learner to Recruit Reinforcement

Another way to "wake up" a potentially powerful but dormant natural contingency of reinforcement is to teach the learner to recruit reinforcement from significant others. For example, Seymour and Stokes (1976) taught delinquent girls to work more productively in the vocational training area of a residential institution. However, observation showed that staff at the institution gave the girls no praise or positive interaction regardless of the quality of their work. The much-needed natural community of reinforcement to ensure the generalization of the girls' improved work behaviors was not functioning. To get around

this difficulty, the experimenters trained the girls to use a simple response that called the attention of staff members to their work. With this strategy, staff praise for good work increased. Thus, teaching the girls an additional response that could be used to recruit reinforcement enabled the target behavior to contact natural reinforcers that would extend and maintain the desired behavior change.

Students of various ages and abilities have learned to recruit teacher and peer attention in classroom and community settings: preschoolers with developmental delays for completing preacademic tasks and staying on task during transitions (Connell, Carta, & Baer, 1993; Stokes, Fowler, & Baer, 1978); and elementary and secondary students with learning disabilities (Alber, Heward, & Hippler, 1999; Wolford, Alber, & Heward, 2001), behavioral disorders (Alber, Anderson, Martin, & Moore, 2004; Morgan, Young, & Goldstein, 1983), and cognitive disabilities (Mank & Horner, 1987; Rouse, Everhart-Sherwood, & Alber-Morgan, 2014) who need help completing various academic and social tasks and improving vocational skills. A study by Craft, Alber, and Heward (1998) provides a representative example of this research.

Craft and colleagues (1998) assessed the effects of recruitment training on academic assignments for which students recruited teacher attention. Four elementary students were trained by their special education teacher (the first author) when, how, and how often to recruit teacher attention in the general education classroom. Training consisted of modeling, role playing, error correction, and praise in the special education classroom. The students were taught to show their work to the teacher or ask for help 2 to 3 times per session, and to use appropriate statements such as "How am I doing?" or "Does this look right?"

Data on the frequency of student recruiting and teacher praise statements were collected during a daily 20-min homeroom period in a general education classroom. During this period, the general education students completed a variety of independent seatwork tasks (reading, language arts, math) assigned by the general education teacher, while the four special education students completed spelling worksheets assigned by the special education teacher, an arrangement that had been established prior to the experiment. If students needed help with their assignments during homeroom, they took their work to the teacher's desk and asked for help.

The effects of recruitment training on the children's frequency of recruiting and on the number of praise statements they received from the classroom teacher are shown in Figure 30.10. Recruiting across students increased from a mean rate of 0.01 to 0.8 recruiting response per 20-min session during baseline to a mean rate of 1.8 to 2.7 after training. Teacher praise statements received by the students increased from a mean rate of 0.1 to 0.8 praise statement per session during baseline to a mean rate of 1.0 to 1.7 after training. The ultimate meaning and outcome of the intervention was in the increased amount and accuracy of academic work by each student (see Figure 6.8).

For a review of research on recruiting and suggestions for teaching children to recruit reinforcement from significant others, see Alber and Heward (2000). Box 30.3: "Look, Teacher, I'm All Finished!" provides suggestions for teaching students to recruit teacher attention.

Mediate Generalization

Another strategy for promoting generalized behavior change is to arrange for some thing or person to act as a medium that ensures the transfer of the target behavior from the instructional setting to the generalization setting. Two tactics for implementing this strategy are contriving a mediating stimulus and teaching the learner to mediate her own generalization through self-management.

Contrive a Mediating Stimulus

One tactic for mediating generalization is to bring the target behavior under the control of a stimulus in the instructional setting that will function in the generalization setting to reliably prompt or aid the learner's performance of the target behavior. The stimulus selected for this important role may exist already in the generalization environment, or it may be a new stimulus added to the instructional program that subsequently goes with the learner to the generalization setting. Whether it is a naturally existing component of the generalization setting or an added element to the instructional setting, to effectively mediate generalization, a **contrived mediating stimulus** must be (a) made functional for the target behavior during instruction and (b) easily transported to the generalization setting (Baer, 1999). The mediating stimulus is *functional* for the learner if it reliably prompts or aids the learner in performing the target behavior; the mediating stimulus is *transportable* if it easily goes with the learner to all important generalization settings.

Naturally existing features of generalization settings that are used as contrived mediating stimuli may be physical objects or people. Van den Pol and colleagues (1981) used paper napkins, a common feature of any fast-food restaurant, as a contrived mediating stimulus. They taught students that a paper napkin was the only place to put food. In this way, the researchers eliminated the added challenge and difficulty of teaching the students to discriminate clean tables from dirty tables, to sit only at clean tables, or to wipe dirty tables, and then programming the generalization and maintenance of those behaviors. By contriving the special use of the napkin, only one response had to be trained, and napkins then served as a mediating stimulus for that behavior.

Cariveau and Kodak (2017) contrived a common stimulus to promote the generalization and maintenance of academic engagement by second-graders during small-group writing and writing lessons. At the start of each lesson during intervention, the instructor placed in front of the students a yellow paper with a numeral written on it representing a performance goal (e.g., 8 indicated a goal of academic engagement for 80% of the session). At the end of each lesson, if a randomly selected student met the performance criterion, the entire group earned a reward. After a reversal design demonstrated a functional relation between the intervention and increased academic engagement, the researchers implemented a maintenance condition that incorporated the common stimulus. "The instructor placed the yellow card with a goal written on it (i.e., the common stimuli) in view of all participants and said, 'This is your goal for today.' At the end of each session, the instructor did not provide information about whether the participants met their goal and did not deliver reinforcement." (p. 126)

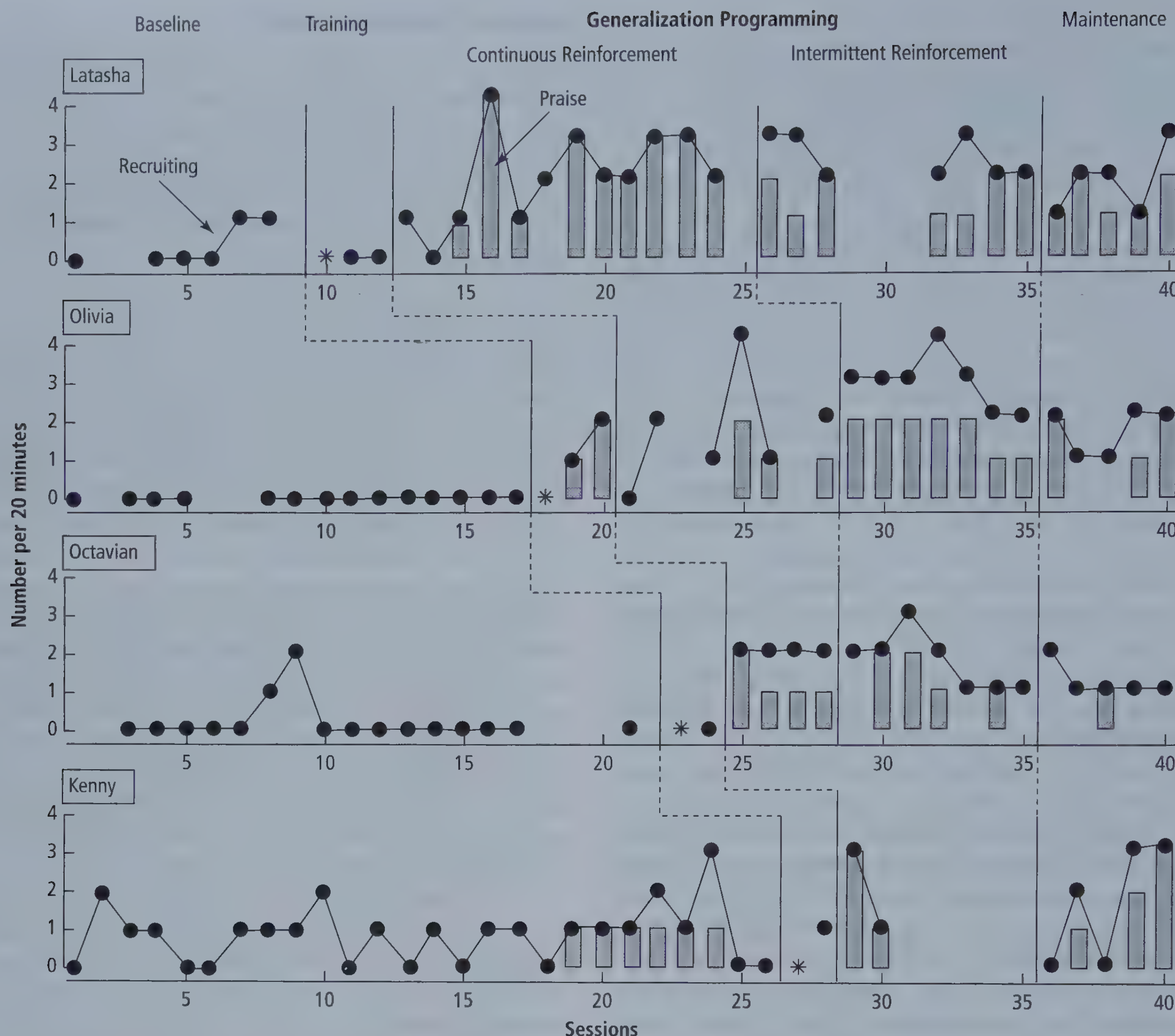


Figure 30.10 Number of recruiting responses (data points) and teacher praise statement (bars) per 20-min session. Target recruiting rate was two to three responses per session. Asterisks indicate when each student was trained in the resource room.

From "Teaching Elementary Students with Developmental Disabilities to Recruit Teacher Attention in a General Education Classroom: Effects on Teacher Praise and Academic Productivity" by M. A. Craft, S. R. Alber, and W. L. Heward, 1998, *Journal of Applied Behavior Analysis, 31,* p. 407. Reproduced with permission of John Wiley & Sons, Inc.

BOX 30.3

"Look, Teacher, I'm All Finished!" Teaching Students to Recruit Teacher Attention

Classrooms are extremely busy places, and even the most conscientious teachers can easily overlook students' important academic and social behaviors. Research shows that teachers are more likely to pay attention to a disruptive student than to one who is working quietly and productively (Walker, 1997). It can be hard for teachers to recognize students who need help, especially low-achieving students who are less likely to ask for help (Newman & Golding, 1990).

Although teachers in general education classrooms are expected to adapt instruction to serve students with disabilities, this is not always the case. Most secondary teachers interviewed by Schumm and colleagues (1995) believed that students with disabilities should take responsibility for obtaining the help they need. Thus, knowing how to politely recruit teacher attention and assistance can help students with disabilities function more independently and actively influence the quality of instruction they receive.

Who Should Be Taught to Recruit?

Although most students would probably benefit from learning to recruit teacher praise and feedback, here are some ideal candidates for recruitment training:

Withdrawn Willamena. Willamena seldom asks a teacher anything. Because she is so quiet and well behaved, her teachers sometimes forget she's in the room.

In-a-Hurry Harry. Harry is usually half-done with a task before his teacher finishes explaining it. Racing through his work allows him to be the first to turn it in. But his work is often incomplete and filled with errors, so he doesn't hear much praise from his teacher. Harry would benefit from recruitment training that includes self-checking and self-correction.

Shouting Shelly. Shelly has just finished her work, and she wants her teacher to look at it—right now! But Shelly doesn't raise her hand. She gets her teacher's attention—and disrupts most of her classmates—by shouting across the room. Shelly should be taught appropriate ways to solicit teacher attention.

Pestering Pete. Pete always raises his hand, waits quietly for his teacher to come to his desk, and then politely asks, "Have I done this right?" But he repeats this routine a dozen or more times in a 20-min period, and his teachers find it annoying. Positive teacher attention often turns into reprimands. Recruitment training for Pete will teach him to limit the number of times he cues his teachers for attention.

How to Get Started

1. **Identify target behaviors.** Students should recruit teacher attention for target behaviors that are valued and therefore likely to be reinforced, such as writing neatly and legibly, working accurately, completing assigned work, cleaning up at transitions, and making contributions when working in a cooperative group.
2. **Teach self-assessment.** Students should self-assess their work before recruiting teacher attention (e.g., Sue asks herself, "Is my work complete?"). After the student can reliably distinguish between complete and incomplete work samples, she can learn how to check the accuracy of her work with answer keys or checklists of the steps or components of the academic skill, or how to spot-check two or three items before asking the teacher to look at it.
3. **Teach appropriate recruiting.** Teach students when, how, and how often to recruit and how to respond to the teacher after receiving attention.
   - *When?* Students should signal for teacher attention after they have completed and self-checked a substantial part of their work. Students should also be taught when not to try to get their teacher's attention (e.g., when the teacher is working with another student, talking to another adult, or taking the lunch count).
   - *How?* The traditional hand raise should be part of every student's recruiting repertoire. Other methods of gaining attention should be taught depending on teacher preferences and routines in the general education classroom (e.g., have students signal they need help by standing up a small flag on their desks; expect students to bring their work to the teacher's desk for help and feedback).
   - *How often?* While helping Withdrawn Willamena learn to seek teacher attention, don't turn her into a Pestering Pete. How often a student should recruit varies across teachers and activities (e.g., independent seatwork, cooperative learning groups, whole-class instruction). Direct observation in the classroom is the best way to establish an optimal rate of recruiting; it is also a good idea to ask the regular classroom teacher when, how, and with what frequency she prefers students to ask for help.
   - *What to say?* Students should be taught several statements that are likely to evoke positive feedback from the teacher (e.g., "Please look at my work." "Did I do a good job?" "How am I doing?"). Keep it simple, but teach the student to vary her verbal cues so she will not sound like a parrot.
   - *How to respond?* Students should respond to their teacher's feedback by establishing eye contact, smiling, and saying "Thank you." Polite appreciation is very reinforcing to teachers.
4. **Role-play the complete sequence.** Begin by providing students with a rationale for recruiting (e.g., the teacher will be happy you did a good job, you will get more work done, your grades might improve). Thinking aloud while modeling is a good way to show the recruiting sequence. While performing each step, say, "Okay, I've finished my work. Now I'm going to check it. Did I put my name on my paper? Yes. Did I do all the problems? Yes. Did I follow all the steps? Yes. Okay, my teacher doesn't look busy right now. I'll raise my hand and wait quietly until she comes to my desk." Have another student pretend to be the regular classroom teacher and come over to you when you have your hand up. Say, "Mr. Patterson, please look at my work." The helper says, "Oh, you did a very nice job." Then smile and say, "Thank you, Mr. Patterson." Role-play with praise and offer corrective feedback until the student correctly performs the entire sequence on several consecutive trials.
5. **Prepare students for alternate responses.** Of course, not every recruiting attempt will result in teacher praise; some recruiting responses may even be followed by criticism (e.g., "This is all wrong. Pay better attention the next time."). Use role playing to prepare students for these possibilities and have them practice polite responses (e.g., "Thank you for helping me with this").
6. **Promote generalization to the regular classroom.** The success of any recruitment training effort depends on the student actually using his or her new skill in the regular classroom.

The maintenance condition was continued until the students' mean level of academic engagement was 20% below the mean level during the second treatment phase for five consecutive sessions. Results for three of the study's six participants are shown in Figure 30.11. Although their experimental design did not include an analysis of the yellow paper's effect on maintenance of academic engagement, Cariveau and Kodak (2017) provided a promising demonstration of the potential power of a contrived common stimulus to mediate generalized behavior change.

When choosing a stimulus to be made common to both teaching and social generalization setting(s), practitioners should consider using people. In addition to being a requisite feature of social settings, people are transportable and important sources of reinforcement for many behaviors. A study by Stokes and Baer (1976) is a good example of the potential influence of the evocative effects of the presence of a person in the generalization setting who had a functional role in the learner's acquisition of the target behavior in the instructional setting. Two preschool children with learning disabilities learned word-recognition skills while working as reciprocal peer tutors for one another. However, neither child showed reliable generalization of those new skills in nontraining settings until the peer with whom he had learned the skills was present in the generalization setting.

Some contrived mediating stimuli serve as more than response prompts; they are prosthetic devices that assist the learner in performing the target behavior. Such devices can be especially useful in promoting generalization and maintenance of complex behaviors and extend response chains by simplifying a complex situation. Three common forms are cue cards, visual activity schedules, and self-operated prompting devices.

Cue Cards. Sprague and Horner (1984) gave the students in their study cue cards to aid them in operating a vending machine without another person's assistance. The cue cards, which had food and drink logos on one side and pictures of quarters paired with prices on the other, not only were used during instruction and generalization probes but also were kept by the students at the end of the program. A follow-up 18 months after the study was completed revealed that five of the six students still carried a cue card and were using vending machines independently.

In a study in which middle school students with learning disabilities learned to recruit peer assistance during cooperative learning group activities in the general education classroom, their special education teacher provided each student with a 3- × 5-in. laminated cue card that listed the recruiting steps (Wolford et al., 2001; see Figure 30.12). The students practiced using the card in the special education class where recruiting training took place. They were encouraged to put the cue cards in their language arts notebooks used in the general education classroom and, if necessary, to look at the card during cooperative learning group activities as a reminder of when and how to recruit.

Visual Activity Schedules. Visual activity schedules use icons, photos, illustrations, and real objects to depict a sequence of events the learner will experience or tasks to complete. Some schedules also include performance criteria and a way to check off completed activities and tasks. Figure 30.13 shows two examples of visual activity schedules. Visual activity schedules can facilitate transitions from one activity to the next, provide structure for down time, help the learner understand what activities are associated with different physical spaces, and promote social interactions and communication skills (Banda, Grimmett, & Hart, 2009; Morrison, Sainato, Benchaaban, & Endo, 2002; Pierce, Spriggs, Gast, & Luscre, 2013). A visual schedule can promote independence and self-determination. Individuals who learn to use visual activity schedules no longer require a teacher, parent, or staff member to tell them when to initiate one activity and when to move on to the next. Beyond a simple prompting tool, a visual schedule can promote independence and self-determination.

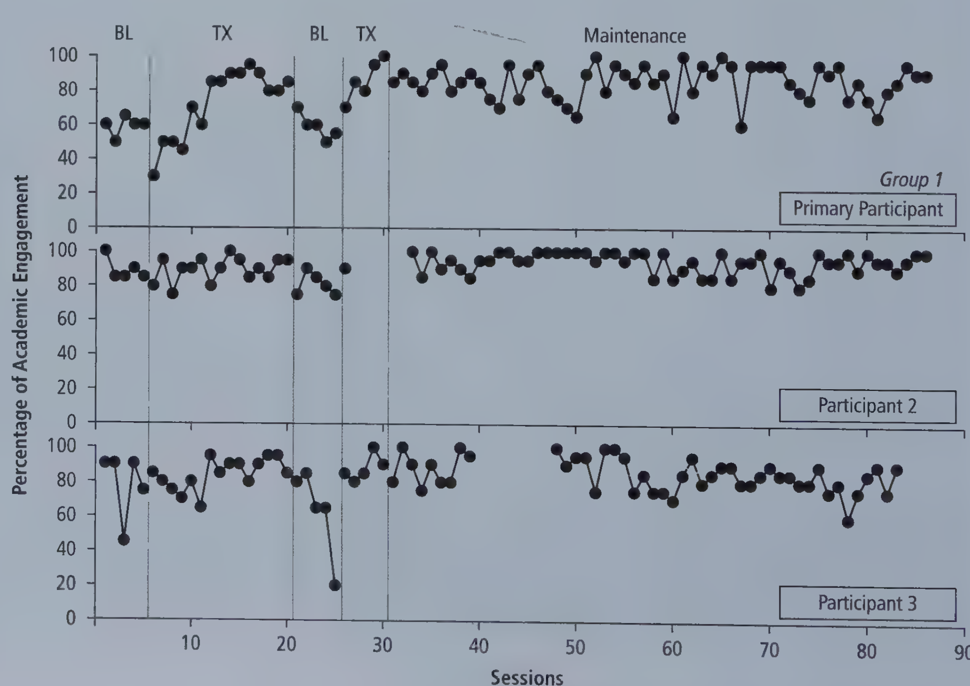


Figure 30.11 Academic engagement by second-grade students during small-group instruction across baseline (BL), treatment (TX), and maintenance conditions. The Primary Participant was the student with the lowest mean level of engagement during baseline. Phase changes were determined by the Primary Participant's data.

From "Programming Randomized Dependent Group Contingency and Common Stimuli to Promote Durable Behavior Change" by T. Cariveau and T. Kodak, 2017, *Journal of Applied Behavior Analysis, 50,* p. 127. Reproduced with permission of John Wiley & Sons, Inc.

Figure 30.12 Middle school students could refer to a laminated card when seeking help from classmates during collaborative learning group activities.

How to Ask for Help

1. Is the teacher talking?
2. Is the student talking to somone else?
3. Get student's attention.
   - say their name quietly
   - say "Excuse me."
   - tap them on shoulder
4. Ask your question clearly.
   - Can you explain what we are supposed to do?
   - How does this look so far?
   - Can you see if I am doing this right?
5. Say "Thank you."

Described in "Teaching Middle School Students with Learning Disabilities to Recruit Peer Assistance During Cooperative Learning Group Activities," by T. Wolford, S. R. Alber, and W. L. Heward, 2001, *Learning Disabilities Research & Practice*, *16*, 161–173.

MacDuff, Krantz, and McClannahan (1993) conducted one of the first behavior analytic experiments evaluating visual activity schedules. They taught four boys with autism, ages 9 to 14, to use visual activity schedules when performing domestic-living skills such as vacuuming and table setting and for leisure activities such as using manipulative toys. The researchers reported that prior to training with photographic activity schedules, the boys required continuous supervision and verbal prompts to complete self-help, housekeeping, and leisure activities. They summarized the results and suggested a possible explanation for the maintenance and generalization as follows.

> When the study ended, all 4 boys were able to display complex home-living and recreational repertoires for an hour, during which time they frequently changed tasks and moved to different areas of their group home without adults' prompts. Photographic activity schedules, . . . became functional discriminative stimuli that promoted sustained engagement after training ceased and fostered generalized responding to new activity sequences and novel leisure materials. (p. 97)

Self-operated Prompting Devices. Numerous studies have shown that learners across a wide range of age and cognitive levels can learn to use mobile audio and video playback devices to independently perform a variety of academic, self-care, domestic, and employment tasks (e.g., Bouck, Satsangi, Muhl, & Bartlett, 2013; Grossi, 1998; Mays & Heflin, 2011; Montgomery, Storey, Post, & Lemley, 2011; Trask-Tyler et al., 1994). The popularity of personal mobile devices such as smartphones and tablets enables a person to listen to or view a series of self-delivered response prompts in a private, normalized manner that does not impose on or bother others (Mechling 2007; Savage, 2014).

Teach Self-management Skills

The most potentially effective approach to mediating generalized behavior changes rests with the one element that is always present in every instructional and generalization setting—the learner herself. Chapter 29 described a variety of self-management tactics that people can use to modify their own behavior. The logic of using self-management to mediate generalized behavior changes goes like this: If the learner can be taught a behavior (not the original target behavior, but another behavior—a controlling response from a self-management perspective) that serves to prompt or reinforce the target behavior in all the relevant settings, at all appropriate times, and in all of its relevant forms, then the generalization of the target behavior is ensured. But as Baer and Fowler (1984) warned:

> Giving a student self-control responses designed to mediate the generalization of some critical behavior changes does not ensure that those mediating responses will indeed be used. They are, after all, just responses: they, too, need generalization and maintenance just as do the behavior changes that they are meant to generalize and maintain. Setting up one behavior to mediate the generalization of another behavior may succeed—but it may also represent a problem in guaranteeing the generalization of two responses, where before we had only the problem of guaranteeing the generalization of one! (p. 149)

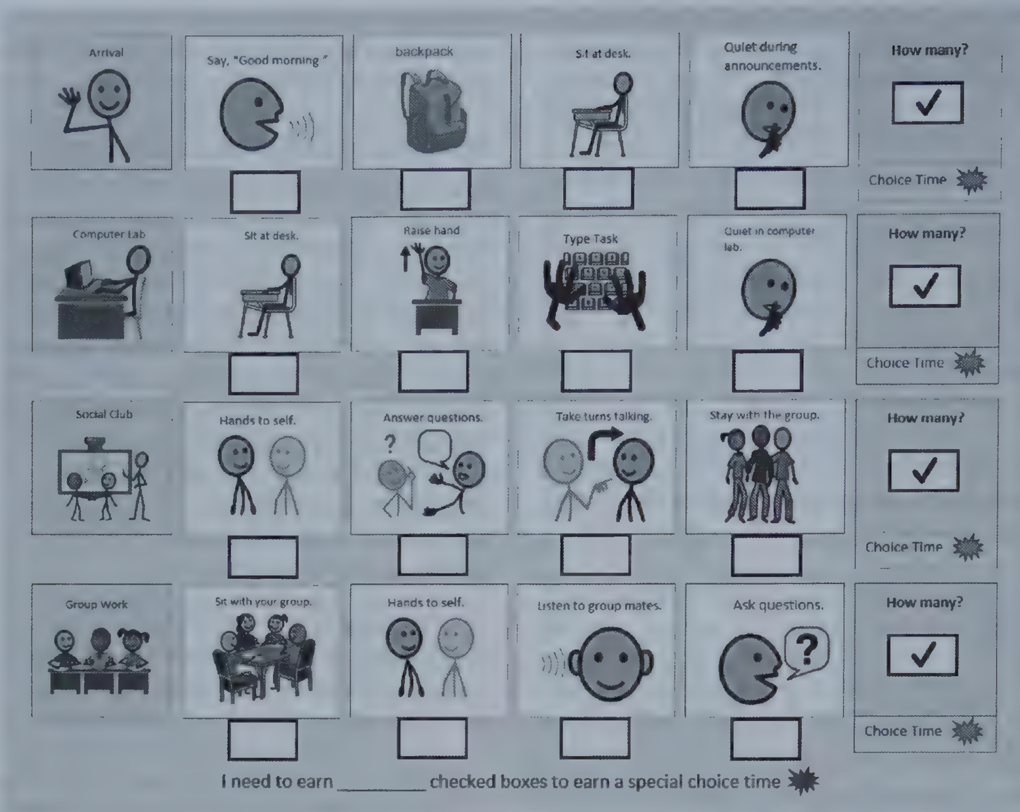
Train to Generalize

> *If generalization is considered as a response itself, then a reinforcement contingency may be placed on it, the same as with any other operant.*
>
> —Stokes and Baer (1977, p. 362)

Training "to generalize" was one of the eight proactive strategies for programming generalized behavior change in Stokes and Baer's (1977) conceptual scheme. The quotation marks around "to generalize" signified that the authors were hypothesizing about the possibilities of treating "to generalize" as an operant response and that they recognized "the preference of behaviorists to consider generalization to be an outcome of behavioral change, rather than as a behavior itself" (p. 363). Although considerable debate surrounds the validity of variable responding being an operant in its own right (e.g., de Souza Barba, 2012; Holth, 2012; Neuringer, 2003, 2004; Peleg, Martin, & Holth, 2017), basic and applied research has demonstrated the pragmatic value of reinforcing varied responding (e.g., Neuringer & Jensen, 2013; Shahan & Chase, 2002). Two tactics that applied behavior analysts have used are reinforcing response variability and instructing the learner to generalize.

Reinforce Response Variability

Response variability can help a person solve problems. A person who can improvise by emitting a variety of responses is more likely to solve problems encountered when a standard response



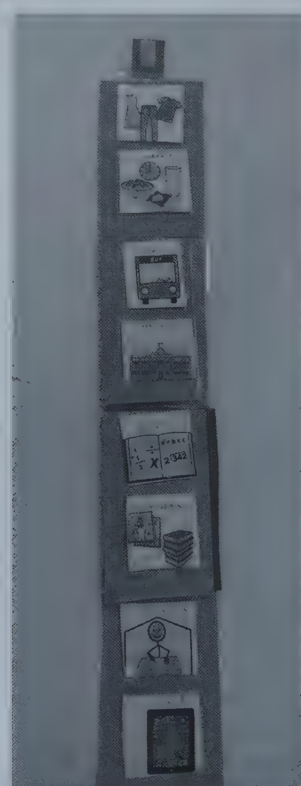


Figure 30.13 The visual activity schedule at the top identifies the sequence of morning activities and criteria for earning the choice of preferred activities during free time. The transportable schedule at the bottom includes changeable icons.

From "Visual Activity Schedules for Students with Autism Spectrum Disorder," by A. Aspen and L. Stack, California Autism Professional Training and Information Network (2017). Used by permission.

form fails to obtain reinforcement (e.g., Arnesen, 2000; Marckel, Neef, & Ferreri, 2006; Miller & Neuringer, 2000; Neuringer, 2004; Shahan & Chase, 2002). Response variability also may result in behavior that is valued because it is novel or creative (e.g., Holman, Goetz, & Baer, 1977; Neuringer, 2003; Pryor, Haag, & O'Reilly, 1969). Response variability may expose a person to sources of reinforcement and contingencies not accessible by more constricted forms of responding. The additional learning that results from contacting those contingencies further expands the person's repertoire.

One direct way to promote desired response generalization is to reinforce response variability when it occurs. The simplest approach entails differential reinforcement of new response forms as Goetz and Baer (1973) did by attending to children when they constructed different blockbuilding forms (see Figure 8.7). Systematic replications of this procedure have ranged from increasing diverse block building by children with autism (Napolitano, Smith, Zarcone, Goodkin, & McAdam, 2010) to varied punching and kicking techniques by martial arts students (Harding, Wacker, Berg, Rick, & Lee, 2004).

The contingency between response variability and reinforcement can be formalized with a lag reinforcement schedule. Reinforcement on a lag schedule is contingent on a response being different in some defined way from the previous response (a Lag 1 schedule) or a specified number of previous responses (Lag 2 or more). Lag reinforcement schedules have produced varied verbal responding by children with disabilities (Contreras & Betz, 2016; Heldt & Schlinger, 2012; see Wiskow & Donaldson's 2016 study discussed in Chapter 13).

Cammilleri and Hanley (2005) used a lag reinforcement contingency to increase varied selections of classroom activities by two typically developing girls, ages 5 and 7, who were selected to participate in the study because they spent the vast majority of their time engaged in activities not systematically designed to result in a particular set of skills within the curriculum. At the beginning of each 60-min session, the children were told they could select any activity and could switch activities at any time. A timer sounded every 5 minutes to prompt activity choices. During baseline, there were no programmed consequences for selecting any of the activities. Intervention consisted of a lag reinforcement schedule in which the first activity selection and each subsequent novel selection were followed by the teacher handing the student a green card that she could exchange later for 2 minutes of teacher attention (resulting in a Lag 12 contingency, which was reset if all 12 activities were selected within a session).

Activity selections during baseline showed little variability, with both girls showing strong preferences for stackable blocks (Figure 30.14 shows results for one of the girls). When the lag contingency was introduced, both girls immediately selected and engaged in more diverse activities. The researchers noted that "an indirect but important outcome of this shift in time allocation was a marked increase in the number of academic units completed" (p. 115).

Instruct the Learner to Generalize

The simplest and least expensive of all tactics for promoting generalized behavior change is to "tell the subject about the possibility of generalization and then ask for it" (Stokes & Baer, 1977, p. 363). For example, Ninness, Fuerst, and Rutherford (1991) explicitly told three middle school students with emotional disturbance to use self-management procedures they had learned in the classroom to self-assess and self-record their behavior while walking from the lunchroom to the classroom. Hughes and colleagues (1995) used a similar procedure to promote generalization: "At the close of each training session, peer teachers reminded participants to self-instruct when they wished to talk to someone" (p. 207). Likewise, at the conclusion of each training session conducted in the special education classroom on how to recruit assistance from peers during cooperative learning groups, Wolford and colleagues (2001) prompted middle school students with learning disabilities to recruit peer assistance at least 2 times but not more than 4 times during each cooperative learning group in the language arts classroom.

To the extent that generalizations occur and are themselves generalized, a person might then become skilled at generalizing newly acquired skills or, in the words of Stokes and Baer (1977), become a "generalized generalizer."

MODIFYING AND TERMINATING SUCCESSFUL INTERVENTIONS

With most successful behavior change programs it is impossible, impractical, or undesirable to continue the intervention indefinitely. Withdrawal of a successful intervention should be carried out in a systematic fashion, guided by the learner's performance of the target behavior in the most important generalization settings. Gradually shifting from the contrived conditions of the intervention to the typical, everyday environment will increase the likelihood that the learner will maintain the new behavior patterns. When deciding how soon and how swiftly to withdraw intervention components, practitioners should consider factors such as the complexity of the intervention, the ease or speed with which behavior changed, and the availability of naturally existing contingencies of reinforcement for the new behavior.

This shift from intervention conditions to the postintervention environment can be made by modifying one or more of the following components, each representing one part of the three-term contingency:

- Antecedents, prompts, or cue-related stimuli
- Task requirements and criteria
- Consequences or reinforcement variables

Although the order in which intervention components are withdrawn may make little or no difference, in most programs it is probably best to make all task-related requirements as similar as possible to those of the postintervention environment before withdrawing significant antecedent or consequence components of the intervention. In this way, the learner will be emitting the target behavior at the same level of performance that will be required after the complete intervention has been withdrawn.

A behavior change program carried out many years ago by a graduate student in one of our classes illustrates how the components of a program can be gradually and systematically withdrawn. A man with developmental disabilities took an inordinate amount of time to get dressed each morning (40 to 70 minutes during baseline), even though he possessed the skills needed to dress himself. Intervention began with a construction paper clock hung by his bed with the hands set to indicate the time by which he had to be fully dressed to receive reinforcement. Although the man could not tell time, he could discriminate whether the position of the hands on the real clock nearby matched those on his paper clock. Two task-related intervention elements were introduced to increase the likelihood of initial success. First, he was

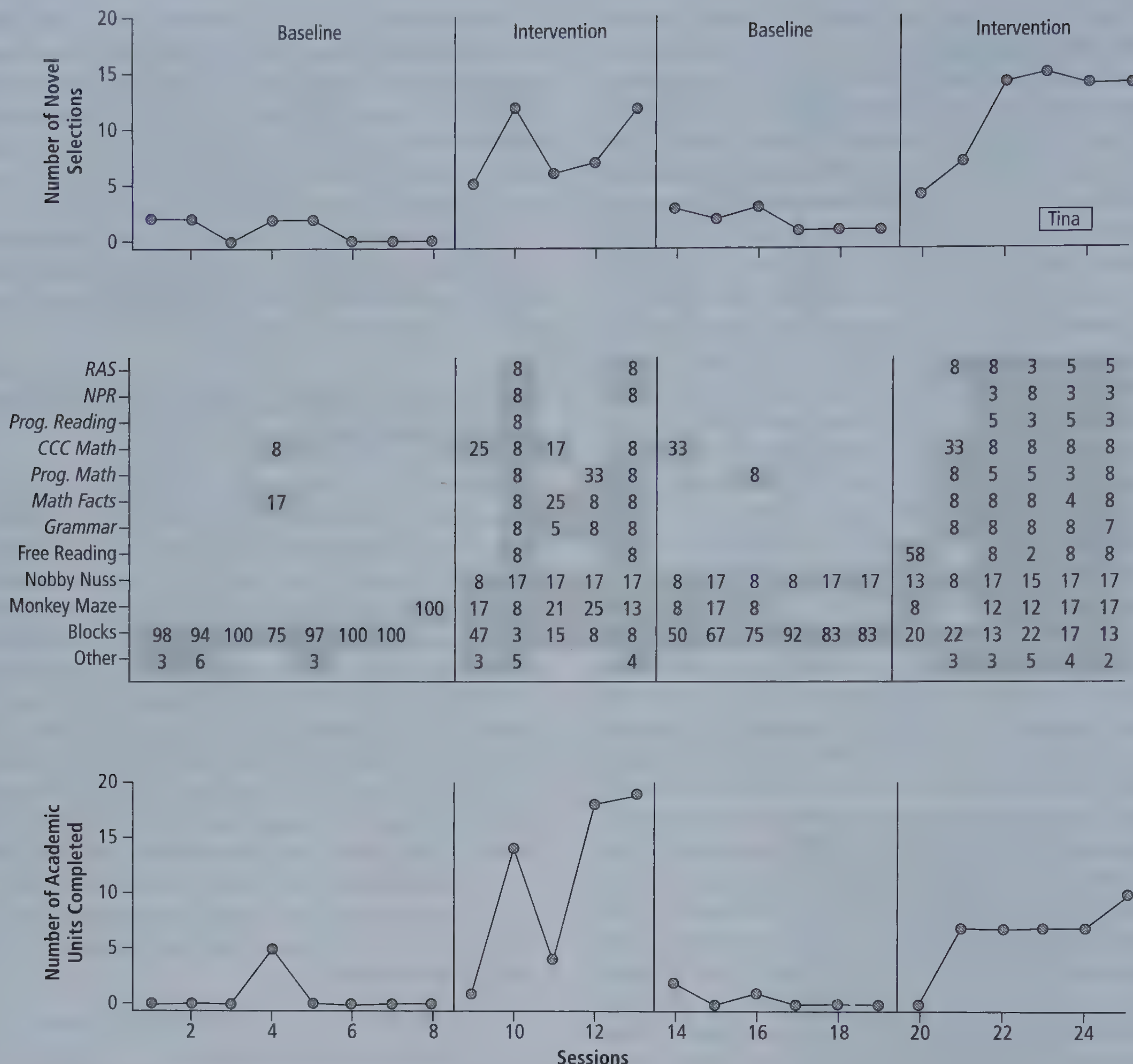


Figure 30.14 Number of novel activity selections (top panel), percentage of intervals of engagement in programmed (italicized) and nonprogrammed activities (middle panel; shaded cells indicate activities for which there was some engagement), and number of academic units completed (bottom panel).

From "Use of a Lag Differential Reinforcement Contingency to Increase Varied Selections of Classroom Activities" by A. P. Cammilleri and G. P. Hanley, 2005, *Journal of Applied Behavior Analysis, 38,* p. 114. Reproduced with permission of John Wiley & Sons, Inc.

given fewer and easier clothes to put on each morning (e.g., no belt, slip-on loafers instead of shoes with laces). Second, based on his baseline performance, he was initially given 30 minutes to dress himself, even though the objective of the program was for him to be completely dressed within 10 minutes. An edible reinforcer paired with verbal praise was used first on a continuous schedule of reinforcement. Figure 30.15 shows how each aspect of the intervention (antecedent, behavior, and consequence) was modified and eventually withdrawn completely, so that by the program's end the man was dressing himself completely within 10 minutes without the aid of extra clocks or charts or contrived reinforcement other than a naturally existing schedule of intermittent praise from staff members.

Rusch and Kazdin (1981) described a method for systematically withdrawing intervention components while simultaneously assessing response maintenance that they called "partial-sequential withdrawal." Martella, Leonard, Marchand-Martella, and Agran (1993) used a partial-sequential withdrawal of various components of a self-monitoring intervention they had implemented to help Brad, a 12-year-old student with mild cognitive disabilities, reduce the number of negative statements (e.g., "I hate this @#!%ing calculator," "Math is crappy") he made during classroom activities. The self-management intervention consisted of Brad (a) self-recording negative statements on a form during two class periods, (b) charting the number on a graph, (c) receiving his choice from a menu of "small" reinforcers (items costing 25 cents or less), and (d) being allowed to choose a "large" reinforcer (more than 25 cents) when his self-recorded data were in agreement with the trainer and at or below a gradually reducing criterion level for four consecutive sessions. After the rate of Brad's negative statements was reduced, the researchers began

Antecedents (cues, prompts)	Behavior (task criteria/ modifications)	Consequences (reinforcers, punishers)
A: Mock clock showing when must be dressed	A: Fewer clothes	A: Edible reinforcer
B: No clock	B: Full set of clothes	B: Token reinforcer
	C: Time gradually reduced	C: Chart (self-record)
	D: Criterion time limit	D: Praise on intermittent schedule
	E: Variety of clothes	

Phase	Antecedents	Behavior	Consequences
1	A	A	A
2	A	A	B
3	A	B	B
4	A	B/C	C
5	A	B/D	C
6	B	B/D	C
7	B	D/E	C
8	B	D/E	D

Figure 30.15 Modification and withdrawal of intervention components to facilitate maintenance and generalization of independent morning dressing by an adult with developmental disabilities.

a four-phase partial-sequential withdrawal of the intervention. In the first phase, charting and earning of "large" reinforcers were withdrawn; in the second phase, Brad was required to have zero negative statements in both periods to receive a daily "small" reinforcer; in the third phase, Brad used the same self-monitoring form for both class periods instead of one form for each period, as before, and the small reinforcer was no longer provided; and in the fourth phase (the follow-up condition), all components of the intervention were withdrawn except for the self-monitoring form without a criterion number highlighted. Brad's negative statements remained low throughout the gradual and partial withdrawal of the intervention (see Figure 30.16).

Another approach to terminating an effective intervention is gradually thinning the schedule of days or sessions in which the full intervention operates. For example, after Falkenberg and Barbetta (2013) determined that homework completion and accuracy by four elementary students had increased and remained high for several weeks as a function of an intervention consisting of self-monitoring and a brief daily conference with their special education teacher to review their self-monitoring sheets, the researchers moved to Fading I and II conditions. In Fading I, the students continued to self-monitor their homework performance daily but met with their teacher only twice per week; in Fading II, self-monitoring continued but conference with the teacher occurred just once per week. Students' homework completion accuracy remained high throughout the two fading conditions and a final maintenance phase in which the students were no longer prompted to self-monitor and no teacher conferences took place.

A word of caution is in order regarding the termination of successful behavior change programs. Achieving socially significant improvements in behavior is a defining purpose of applied behavior analysis. Additionally, those improved behaviors should be maintained and should show generalization to other relevant settings and behaviors. In most instances, achieving optimal generalization of the behavior change will require most, if not all, of the intervention components to be withdrawn. However, practitioners, parents, and others responsible for helping children learn important behaviors are sometimes more concerned with whether and with how a potentially effective intervention will eventually be withdrawn than they are with whether it will produce the needed behavior change. Considering how a proposed intervention will lend itself to eventual withdrawal or blending into the natural environment is important and consistent with everything recommended throughout this book. And clearly, when the choice is between two or more interventions of potentially equal effectiveness, the intervention that is most similar to the natural environment and the easiest to withdraw and terminate should be given first priority. However, an important behavior change should not be forsaken because complete withdrawal of the intervention required to achieve it is considered unlikely. Some level of intervention may always be required to maintain certain behaviors.

GUIDING PRINCIPLES FOR PROMOTING GENERALIZED OUTCOMES

Regardless of the specific tactics selected and applied, we believe that practitioners' efforts to promote generalized behavior change will be enhanced by adherence to five guiding principles:

1. Minimize the need for generalization as much as possible.
2. Conduct generalization probes before, during, and after instruction.
3. Involve significant others whenever possible.
4. Promote generalization with the least intrusive, least costly tactics possible.
5. Contrive intervention tactics as needed to achieve important generalized outcomes.

Minimize the Need for Generalization

Practitioners should reduce the need for generalization to untaught skills, settings, and situations as much as possible. Doing so requires thoughtful, systematic assessment of what behavior changes are most important. Practitioners should

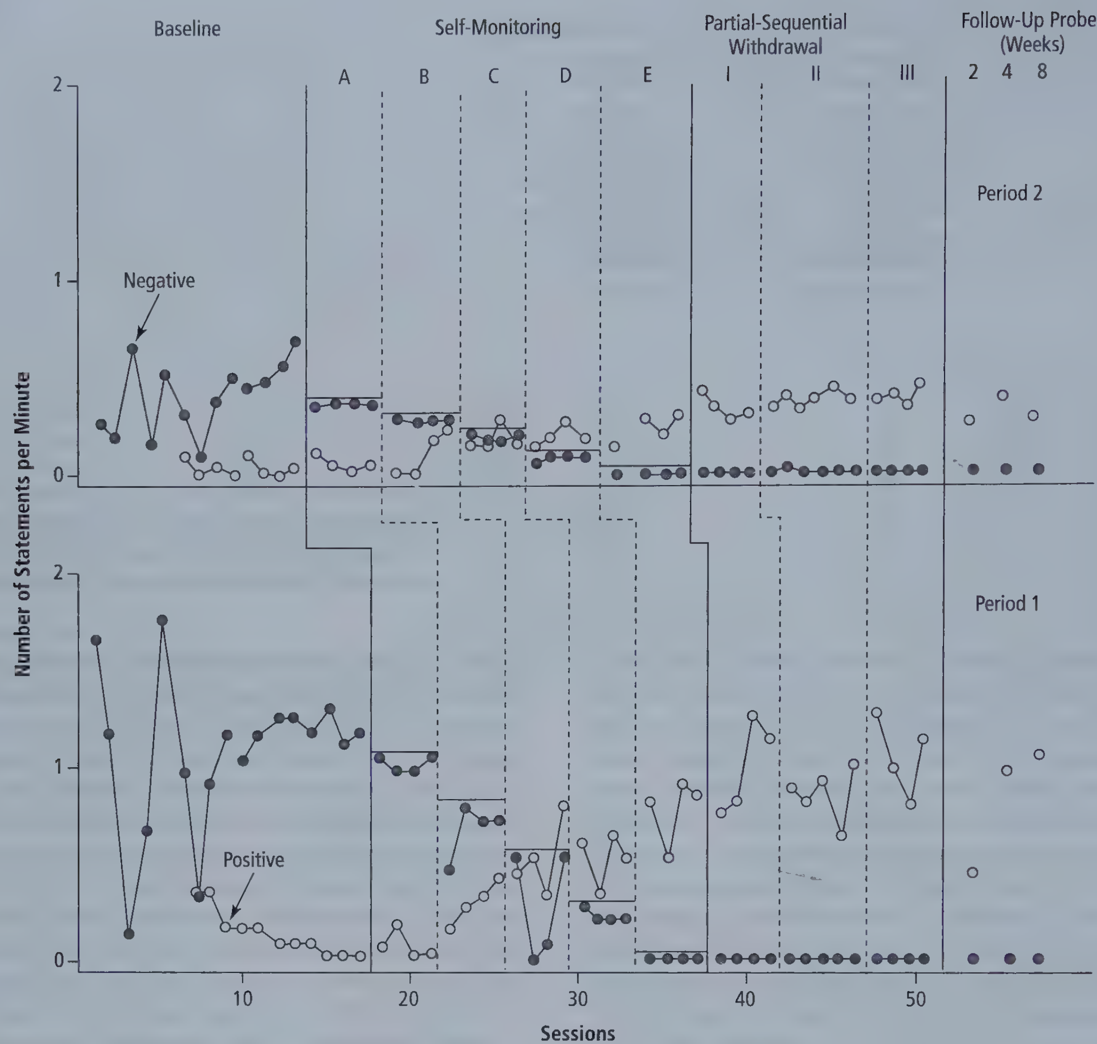


Figure 30.16 Number of negative (solid data points) and positive (open data points) statements in two class periods by a 12-year-old boy with disabilities during baseline, self-monitoring, and partial-sequential withdrawal conditions. Horizontal lines during the self-monitoring condition show changing criteria for teacher-delivered reinforcement.

"Self-Monitoring Negative Statements" by R. Martella, I. J. Leonard, N. E. Marchand-Martella, and M. Agran, 1993, *Journal of Behavioral Education, 3,* p. 84. Copyright 1993 by Human Sciences Press. Reprinted by permission.

prioritize the knowledge and skills that will most often be required of the learner and the settings and situations in which the learner will most often benefit from using those skills. In addition to the environment(s) in which the learner is currently functioning, practitioners should consider the environments in which the learner will function in the immediate future and later in life.

The most critical behavior changes should not be relegated to the not-for-certain technology of generalization. The most important skill-setting-stimulus combinations should always be taught directly and, when possible, taught first. For example, a training program to teach a young adult with disabilities to ride the public bus system should use the bus routes the person will take most often (e.g., to and from her home, school, job site, and community recreation center) as teaching examples. Instead of providing direct instruction on routes to destinations the learner may visit only occasionally after training, use those routes as generalization probes. Achieving a high level of response maintenance on the trained routes will still be a challenge.

Probe for Generalization Before, During, and After Instruction

Generalization probes should be conducted prior to, during, and after instruction.

Probe Before Instruction

Probes conducted prior to instruction may reveal that the learner already performs some or all of the needed behaviors in the generalization setting, thereby reducing the scope of the teaching task. It is a mistake to assume that because a learner does not perform a particular behavior in the instructional setting, she does not or cannot do it in the generalization setting(s). Preinstruction probe data are the only objective basis for knowing that the learner's performance of the target behavior *after instruction* is, in fact, a generalized outcome.

Additionally, probes before instruction enable observation of the contingencies that operate in the generalization setting.

Knowledge of such information may contribute to a more effective treatment or instruction.

Probe During Instruction

Probes during instruction reveal if and when generalization has occurred and if and when instruction can be terminated or shifted in focus from acquisition to maintenance. A teacher who finds that a student can solve untaught examples of a particular type of algebraic equation after receiving instruction on just a few examples, and shifts instruction to the next type of equation in the curriculum, will cover more of the algebra program effectively than will a teacher who continues to present additional examples.

Probes during instruction also show whether generalization is not occurring, thereby indicating that a change in instructional strategies is needed. When Sprague and Horner (1984) observed that one student's poor performance on a generalization probe was caused by a ritualistic pattern of inserting coins, they incorporated repeated practice on the coin insertion step during a training session and the student's performance on the generalization task improved greatly.

In a study on teaching children with autism to respond with empathy to others, Sivaraman (2017) randomly interspersed generalization probes during all treatment sessions. Probes were presented in the same manner as during baseline. The nontraining discriminative stimuli presented during each probe trial belonged to the same antecedent emotion category (frustration, sadness, or happiness) that was being trained. To assess the extent to which tackling by high school football players generalized from practice with a dummy bag at walking speed to tackling a live ball carrier, Harrison and Pyles (2013) conducted probe sessions of four-tackle generalization probes at each stage of a shaping program (see Figure 22.4).

Generalization probes can often be made more efficient by contriving opportunities for the learner to use her new knowledge or skill. For example, instead of waiting for (and perhaps missing) naturally occurring opportunities for the learner to use her new conversational skills in the generalization environment, a practitioner could enlist the assistance of a "confederate" peer to approach the learner. Ninness et al. (1991) used contrived generalization probes in which students were provoked or distracted to test the extent to which they used a set of self-management skills.

Contrived generalization probes can also be used as the primary measure of acquisition and generalization. For example, Miltenberger and colleagues (2005) contrived opportunities to measure and teach children's use of gun safety skills by placing a gun in home and school locations where the children would find it. If a child did not execute the target safety skills when finding a gun (i.e., do not touch it, get away from the gun, and tell an adult), the trainer entered the room and conducted an *in situ* training session, asking the child what he should have done and rehearsing the entire sequence 5 times.

Probe After Instruction

Probes after instruction has ended reveal the extent of response maintenance. The question of how long probes should be conducted after instruction has ended should be answered by factors such as the severity of the target behavior, the importance of the behavior change to the person's quality of life, the strength and consistency of the response maintenance obtained by probes to date, and so forth. The need for long-term assessment of maintenance is especially critical for severe behavior problems. In some cases, maintenance probes over many months and several years may be indicated (e.g., Derby et al., 1997; Wagman, Miltenberger, & Woods, 1995). Reid, Parsons, and Jensen (2017) reported that after intensive staff training at a center for adolescents and adults with severe disabilities brought about desired performance (increased participant involvement in functional tasks), supervisors observed staff and provided performance feedback approximately once per week for 12 weeks and thereafter monthly. This program of ongoing monitoring and feedback maintained high levels of staff performance over follow-up observations spanning 30 years.

If conducting systematic generalization and maintenance probes seems too difficult or too contrived, the practitioner should consider the relative importance of the target behavior to the student or client. If a behavior is important to target for intervention, then assessing the generalized outcomes of that intervention is worth whatever effort that assessment requires.

Involve Significant Others

> *All persons are potential teachers of all sorts of behavior changes. Just because you are designated as "teacher" or "behavior analyst" does not mean that you have an exclusive franchise on the ability to make deliberate behavior changes. In fact, there is no possibility of such a franchise. Everyone in contact contributes to everyone else's behavior, both to its changes and its maintenance.*
>
> —Donald M. Baer (1999, p. 12)

Teaching for generalized outcomes is a big job, and practitioners should try to get as much help as they can. People are almost always around where and when important behaviors need to be prompted and reinforced; and when social behaviors are targeted, people are there by definition.

In virtually every behavior change program, people other than the participant and the behavior analyst are involved, and their cooperation is crucial to the success of the program. Foxx (1996) stated that in programming successful behavior change interventions, "10% is knowing what to do; 90% is getting people to do it. . . . Many programs are unsuccessful because these percentages have been reversed" (p. 230). Although Foxx was referring to the challenge and importance of getting staff to implement programs with the consistency and fidelity needed for success, the same holds for the involvement of significant others.

Baer (1999) suggested identifying others who will or may be involved in a behavior change program as active supporters or tolerators. An active supporter is someone naturally present in the generalization setting who helps promote the generalization and maintenance of the target behavior by doing specific things. Active supporters help facilitate the desired generalized outcomes by arranging opportunities for the learner to use or practice the new skill, giving cues and response prompts for the behavior, and providing reinforcement for performance of the target behavior.

Active supporters for a behavior change program to teach independent eating to a child with severe disabilities might include one or two key people in the school's cafeteria, a volunteer or aide who works regularly with the child, the child's parents, and an older sibling. These people are vital parts of the teaching team. If optimal generalization is to occur, the active supporters must see to it that the learner has many opportunities to use the new skill in the generalization environments they share, and the natural reinforcers they control in those environments (e.g., praise, touch, smiles, companionship) must be used as consequences for the target behavior. It is not necessary, and probably not desirable, to limit an active supporters list to school staff, family members, and peers. They may not be regularly available in all the environments and situations in which generalization is desired.

A tolerator is someone in the generalization setting who agrees not to behave in ways that would impede the generalization plan. Tolerators should be informed that the learner will be using a new skill in the generalization environment and asked to be patient. In an independent eating program, a list of tolerators would likely include some members of the child's family, school cafeteria staff, and peers. Beyond the home and school environments, the behavior analyst should consider the possible role of the general public with whom the child may find herself sharing a table or dining area in a public restaurant. The learner's initial sloppiness and slowness (she may always be sloppier and slower than most people) as she makes the first attempts beyond the familiar contingencies of home and school might result in various responses from strangers that could punish the new eating skills. Being stared at, laughed at, talked about, told to hurry up, or even offered assistance could reduce the possibility of generalization. Certainly, the behavior analyst can inform various school staff and family members of the ongoing eating program and request that they not interfere with the child's attempts to eat independently. But the public at large is another issue. It is impossible to inform everyone of the program. However, by considering the types of intolerant behaviors that the learner may encounter in the generalization setting, the teaching program can be constructed to include practice under such intolerant conditions. Instructional trials might be contrived to reinforce the learner for ignoring rude remarks and continuing to eat independently.

Use the Least Intrusive, Least Costly Tactics Possible

Behavior analysts should use less intrusive and less costly tactics to promote generalization before using more intrusive and more costly tactics. As noted previously, simply reminding students to use their new skills in the generalization setting is the easiest and least expensive of all methods that might promote generalization. Although practitioners should never assume that telling the learner to generalize will produce the desired outcomes, neither should they fail to include such a simple and cost-free tactic. Similarly, incorporating some of the most relevant features of the generalization setting into the instructional setting (i.e., programming common stimuli) often helps produce the needed generalization and is less costly than conducting instruction in the natural setting (e.g., Neef, Lensbower et al., 1990; van den Pol et al., 1981).

Not only is using less costly tactics good conservation of the limited resources available for teaching, but also less intrusive interventions with fewer moving parts are easier to withdraw. Systematic generalization probes will determine whether generalization has occurred and whether more elaborate and intrusive intervention and supports are needed.

Contrive Intervention Tactics as Needed to Achieve Important Generalized Outcomes

A practitioner should not be so concerned about intrusiveness that she fails to implement a potentially effective intervention or procedures that will achieve important outcomes for the learner. Therefore, if necessary, practitioners should disregard the previous guideline and contrive as many instruction and generalization tactics as necessary to enable the learner to generalize and maintain critical knowledge and skills.

Rather than lament the lack of generalization or blame the learner for his inability to show generalized behavior changes, the behavior analyst should work to arrange whatever socially valid contingencies may be needed to extend and maintain the target behavior.

Some Final Words of Wisdom from Don Baer

The most difficult and important challenge facing behavioral practitioners is helping learners achieve generalized change in socially significant behaviors. A behavior change—no matter how important initially—is of little value to the learner if it does not last over time, is not emitted in appropriate settings and situations, or occurs in restricted form when varied topographies are desired.

Research during the past 40 years has developed and advanced the knowledge base of what Stokes and Baer (1977) described as an "implicit technology of generalization" into an increasingly explicit and effective set of strategies and tactics for promoting generalized behavior change. Knowledge of these methods, combined with knowledge of the basic principles and behavior change tactics described throughout this book, provides behavior analysts with a powerful approach for helping people enjoy healthy, happy, and productive lives.

We end this chapter with a dually wise observation by Don Baer (1999), in which he pointed out a fundamental truth about the relation between a person's experience (in this case, the nature of a lesson) and what is learned or not learned from that experience. Like Skinner before him, Baer wisely reminded us not to blame the learner for not behaving as we think he should.

> Learning one aspect of anything never means that you will automatically know the rest of it. Doing something skillfully now never means that you will always do it well. Resisting one temptation consistently never means that you now have character, strength, and discipline. Thus, it is not the learner who is dull, learning disabled, or immature, because all learners are alike in this regard: *no one learns a generalized lesson unless a generalized lesson is taught.* (p. 1) [emphasis in original]

SUMMARY

Generalized Behavior Change: Definitions and Key Concepts

1. Generalized behavior change has taken place if trained behavior occurs at other times or in other places without having to be retrained completely in those times or places, or if functionally related behaviors occur that were not taught directly.
2. Response maintenance refers to the extent to which a learner continues to perform a behavior after a portion or all of the intervention responsible for the behavior's initial appearance in the learner's repertoire has been terminated.
3. Setting/situation generalization refers to the extent to which a learner emits the target behavior in settings or situations that are different from the instructional setting.
4. The instructional setting is the environment where instruction occurs and encompasses all aspects of the environment, planned or unplanned, that may influence the learner's acquisition and generalization of the target behavior.
5. A generalization setting is any place or stimulus situation that differs from the instructional setting in some meaningful way and in which performance of the target behavior is desired.
6. Response generalization refers to the extent to which a learner emits untrained responses that are functionally equivalent to the trained response.
7. Some interventions yield significant and widespread generalized effects across time, settings, and other behaviors; others produce circumscribed changes in behavior with limited endurance and spread.
8. Undesirable setting/situation generalization takes two common forms: overgeneralization, in which the behavior has come under control of a stimulus class that is too broad, and faulty stimulus control, in which the behavior comes under the control of an irrelevant antecedent stimulus.
9. Undesired response generalization occurs when any of a learner's untrained but functionally equivalent responses produce undesirable outcomes.
10. Other types of generalized outcomes (e.g., stimulus equivalence, contingency adduction, and generalization across subjects) do not fit easily into categories of response maintenance, setting/situation generalization, and response generalization.
11. The generalization map is a conceptual framework for combining and categorizing the various types of generalized behavior change.

Planning for Generalized Behavior Change

12. The first step in promoting generalized behavior changes is to select target behaviors that will meet naturally existing contingencies of reinforcement.
13. A naturally existing contingency is any contingency of reinforcement (or punishment) that operates independent of the behavior analyst's or practitioner's efforts, including socially mediated contingencies contrived by other people and already in effect in the relevant setting.
14. A contrived contingency is any contingency of reinforcement (or punishment) designed and implemented by a behavior analyst to achieve the acquisition, maintenance, and/or generalization of a targeted behavior change.
15. Planning for generalization includes identifying all the desired behavior changes and all the environments in which the learner should emit the target behavior(s) after direct training has ceased.
16. Benefits of developing the planning lists include a better understanding of the scope of the teaching task and an opportunity to prioritize the most important behavior changes and settings for direct instruction.

Strategies and Tactics for Promoting Generalized Behavior Change

17. Researchers have developed and advanced what Stokes and Baer (1977) called an "implicit technology of generalization" into an increasingly explicit and effective set of methods for promoting generalized behavior change.
18. The strategy of teaching enough examples requires teaching a subset of all of the possible stimulus and response examples and then assessing the learner's performance on untrained examples.
19. A generalization probe is any measurement of a learner's performance of a target behavior in a setting and/or stimulus situation in which direct training has not been provided.
20. Teaching enough stimulus examples involves teaching the learner to respond correctly to more than one example of an antecedent stimulus and probing for generalization to untaught stimulus examples.
21. As a general rule, the more examples the practitioner uses during instruction, the more likely the learner will be to respond correctly to untrained examples or situations.
22. Having the learner practice a variety of response topographies helps ensure the acquisition of desired response forms and promotes response generalization. Often called multiple-exemplar training, this tactic typically incorporates numerous stimulus examples and response variations.
23. General case analysis is a systematic method for selecting teaching examples that represent the full range of stimulus variations and response requirements in the generalization setting.

24. Negative, or "don't do it," teaching examples help learners identify stimulus situations in which the target behavior should not be performed.

25. Minimum difference negative teaching examples, which share many characteristics with positive teaching examples, help eliminate "generalization errors" due to overgeneralization and faulty stimulus control.

26. The greater the similarity between the instructional setting and the generalization setting, the more likely the target behavior will be emitted in the generalization setting.

27. Programming common stimuli means including in the instructional setting stimulus features typically found in the generalization setting. Practitioners can identify possible stimuli to make common by direct observation in the generalization setting and by asking people who are familiar with the generalization setting.

28. Teaching loosely—randomly varying noncritical aspects of the instructional setting within and across teaching sessions—(a) reduces the likelihood that a single or small group of noncritical stimuli will acquire exclusive control over the target behavior and (b) makes it less likely that the learner's performance will be impeded or "thrown off" by the presence of a "strange" stimulus in the generalization setting.

29. A newly learned behavior may fail to contact an existing contingency of reinforcement because it has not been taught well enough. The solution for this kind of generalization problem is to teach the learner to emit the target behavior at the rate, accuracy, topography, latency, duration, and/or magnitude required by the naturally occurring contingencies of reinforcement.

30. The use of intermittent schedules of reinforcement and delayed rewards can create indiscriminable contingencies, which promote generalized responding by making it difficult for the learner to discriminate whether the next response will produce reinforcement.

31. Behavior traps are powerful contingencies of reinforcement with four defining features: (a) They are "baited" with virtually irresistible reinforcers; (b) only a low-effort response already in the student's repertoire is needed to enter the trap; (c) interrelated contingencies of reinforcement inside the trap motivate the student to acquire, extend, and maintain targeted skills; and (d) they can remain effective for a long time.

32. One way to wake up an existing but inoperative contingency of reinforcement is to ask key people in the generalization setting to attend to and praise the learner's performance of the target behavior.

33. Another tactic for waking up a natural contingency of reinforcement is to teach the learner how to recruit reinforcement in the generalization setting.

34. One tactic for mediating generalization is to bring the target behavior under the control of a contrived stimulus in the instructional setting that will reliably prompt or aid the learner's performance of the target behavior in the generalization setting.

35. Teaching a learner self-management skills with which he can prompt and maintain targeted behavior changes in all relevant settings at all times is the most potentially effective approach to mediating generalized behavior changes.

36. The strategy of training to generalize is predicated on treating "to generalize" as an operant response class that, like any other operant, is selected and maintained by contingencies of reinforcement.

37. One tactic for promoting response generalization is to reinforce response variability. On a lag reinforcement schedule, reinforcement is contingent on a response being different in some defined way from the previous response (a Lag 1 schedule) or a specified number of previous responses (Lag 2 or more).

38. The simplest and least expensive tactic for promoting generalized behavior change is to tell the learner about the usefulness of generalization and then instruct him to do so.

Modifying and Terminating Successful Interventions

39. With most successful behavior change programs it is impossible, impractical, or undesirable to continue the intervention indefinitely.

40. The shift from formal intervention procedures to a normal everyday environment can be accomplished by gradually withdrawing elements constituting the three components of the training program: (a) antecedents, prompts, or cue-related stimuli; (b) task modifications and criteria; and (c) consequences or reinforcement variables.

41. An important behavior change should not go unmade because complete withdrawal of the intervention required to achieve it may never be possible. Some level of intervention may always be required to maintain certain behaviors, in which case attempts must be made to continue necessary programming.

Guiding Principles for Promoting Generalized Outcomes

42. Efforts to promote generalized behavior change will be enhanced by adhering to five guiding principles:
    - Minimize the need for generalization as much as possible.
    - Conduct generalization probes before, during, and after instruction.
    - Involve significant others whenever possible.
    - Promote generalized behavior change with the least intrusive, least costly tactics possible.
    - Contrive intervention tactics as needed to achieve important generalized outcomes.

KEY TERMS

behavior trap
contrived contingency
contrived mediating stimulus
general case analysis
generalization
generalization across subjects
generalization probe
generalization setting
indiscriminable contingency
instructional setting
lag reinforcement schedule
multiple exemplar training
naturally existing contingency
programming common stimuli
response generalization
response maintenance
setting/situation generalization
teaching sufficient examples
teaching loosely

MULTIPLE-CHOICE QUESTIONS

1. This is the extent to which a learner emits the target behavior in a setting or stimulus situation that is different from the instructional setting.
   a. Setting/situation maintenance
   b. Response generalization
   c. Setting/situation generalization
   d. Response maintenance
   Hint: (See "Generalized Behavior Change: Definitions and Key Concepts")

2. This is the extent to which a learner emits untrained responses that are functionally equivalent to the trained target behavior
   a. Setting/situation maintenance
   b. Response generalization
   c. Setting/situation generalization
   d. Response maintenance
   Hint: (See "Generalized Behavior Change: Definitions and Key Concepts")

3. A learner's untrained but functionally equivalent responses result in poor performance. This would be an example of:
   a. Overgeneralization
   b. Faulty stimulus control
   c. Undesirable outcome of setting/situation generalization
   d. Undesirable outcome of response generalization
   Hint: (See "Generalized Behavior Change: Definitions and Key Concepts")

4. When a target behavior comes under the restricted control of an irrelevant antecedent stimulus, it is an example of:
   a. Overgeneralization
   b. Faulty stimulus control
   c. Undesirable outcome of response maintenance
   d. Undesirable outcome of response generalization
   Hint: (See "Generalized Behavior Change: Definitions and Key Concepts")

5. A contingency that operates without social mediation is referred to as:
   a. Natural contingency
   b. Contrived contingency
   c. Unplanned contingency
   d. Generalized contingency
   Hint: (See "Planning for Generalized Behavior Change")

6. Choosing only those behaviors to change that will produce reinforcers in the post-intervention environment is referred to as:
   a. Teaching loosely
   b. Contrived contingency planning
   c. Programming common stimuli
   d. Relevance-of-behavior rule
   Hint: (See "Planning for Generalized Behavior Change")

7. The actual number of examples needed to be taught when teaching sufficient stimulus examples to a learner will depend on:
   a. The learner's opportunities to emit the target behavior under various conditions
   b. Naturally existing contingencies of reinforcement
   c. The learner's history of reinforcement for generalized responding
   d. All of these
   Hint: (See "Strategies and Tactics for Promoting Generalized Behavior Change")

8. When using multiple exemplar training,
   a. Only stimulus variations should be used
   b. Only response variations should be used
   c. Both stimulus and response variations should be used
   d. Neither stimulus nor response variations should be used
   Hint: (See "Strategies and Tactics for Promoting Generalized Behavior Change")

9. Including typical features of the generalization setting into the instructional setting is referred to as:
   a. Teaching loosely
   b. Teaching negative examples
   c. General case analysis
   d. Programming common stimuli
   Hint: (See "Strategies and Tactics for Promoting Generalized Behavior Change")

10. Shifting from intervention to post-intervention conditions can be accomplished by modifying:

a. Antecedents, prompts, or cue-related stimuli

b. Task requirements and criteria

c. Consequences or reinforcement variables

d. All of these

Hint: (See "Modifying and Terminating Successful Interventions")

11. A guiding principle for promoting generalized behavior change includes:

a. Involving significant others whenever possible

b. Promoting generalization with the least intrusive, least costly tactics possible

c. Contriving intervention tactics as needed to achieve important generalized outcomes

d. All of these

Hint: (See "Guiding Principles for Promoting Generalized Outcomes")

ESSAY-TYPE QUESTIONS

1. Name and discuss the three major types of generalized behavior change.

Hint: (See "Generalized Behavior Change: Definitions and Key Concepts")

2. Discuss overgeneralization and faulty stimulus control, and how each can be detrimental to the generalized behavior change for a learner.

Hint: (See "Generalized Behavior Change Is not always desirable")

3. Explain the most important criterion that the text discusses when planning for generalized behavior change.

Hint: (See "Planning for Generalized Behavior Change")

4. There are a number of strategies and tactics that can be utilized to promote generalized behavior change. What are three of these strategies, and how does each one aim to promote generalized behavior change for the learner?

Hint: (See "Strategies and Tactics for Prompting Generalized Behavior Change")

5. When deciding how soon to withdraw an intervention or components of the intervention, what should be considered?

Hint: (See "Modifying and Terminating Successful Interventions")

6. What are the five guiding principles for promoting generalized behavior change?

Hint: (See "Guiding Principles for Promoting Generalized Outcomes")

NOTES

1. Wacker and colleagues (2017) proposed another conception of maintenance: the persistence of treatment effects when treatment is challenged.
2. Because the majority of the examples in this chapter are school based, we use the language of education. To help generalize the examples beyond schooling, read *instruction* as a synonym for *treatment, intervention,* or *therapy; instructional setting* as a synonym for *clinical setting* or *therapy setting;* and *teacher* as a synonym for *practitioner, clinician,* or *therapist.*
3. The unintended maintenance of problem behavior that may result from the manner in which interventions are withdrawn is an undesirable outcome (Ringdahl & St. Peter, 2017). See discussion of resurgence in Chapter 24.
4. Examples of faulty stimulus control caused by flaws in the design of instructional materials, and suggestions for detecting and correcting those flaws, can be found in J. S. Vargas (1984).
5. Teaching a new behavior without developing and implementing a plan to facilitate its maintenance and generalization is so common that Stokes and Baer (1977) called it the "train and hope" approach to generalization.
6. Other terms commonly used for this strategy for promoting generalized behavior change are *training sufficient exemplars, training diversely,* and *multiple-exemplar training.*
7. LaFrance, Tarbox, and Holth (2019) proposed the term *mixed operant instruction* for teaching that features a rapid rotation between consecutive trials that target different operants. Teaching the same target (e.g., water) across verbal operants (e.g., mand, tact, intraverbal) can promote functional interdependence between verbal operants and between speaker and listener repertoires (Fiorile & Greer, 2007; Greer, Stolfi, Chavez-Brown, & Rivera-Valdes, 2005; Nuzzolo-Gomez & Greer, 2004; Olaff, Ona, & Holth, 2017).
8. Holth (2017) identifies two situations in which multiple-exemplar training is ineffective: when (a) there are no physical dimensions along which generalized responding can emerge, or (b) the relation between an antecedent stimulus and an effective response is complex.
9. For a thorough and sophisticated treatment of the selection and sequencing of teaching examples for optimally effective and efficient curriculum design, see Engelmann and Carnine's (1982) *Theory of Instruction: Principles and Applications.* Critchfield and Twyman (2014) and Twyman, Layng, Stikeleather, and Hobbins (2004) also discuss curriculum design from the standpoint of maximizing learning outcomes.
10. Teaching examples used to help students discriminate when not to respond (i.e., S^{Δ}s) are sometimes called *negative examples* and contrasted with positive examples (i.e., S^{D}s). However, in our work in teaching this concept, practitioners have told us that the term *negative teaching example* suggests that the teacher is modeling or showing the learner *how* not to perform the target behavior. Instruction on the desired topography for some behaviors may be aided by providing students with models on how not to perform certain behavior (i.e., negative examples), but the function of "don't do it" examples is to help the learner discriminate antecedent conditions that signal an inappropriate occasion for responding.
11. Coach Hill's effort to simulate the din opposing teams must deal with at "The Shoe" may have helped; his team scored twice. A bit of programming for generalization, however, could not overcome other variables, most notably the home team's overwhelming prowess. Final score: Buckeyes 43, Bulldogs 10.

PART 13

Ethics

In Chapter 31, Tom Freeman, Linda LeBlanc, and Jose Martinez-Diaz help behavior analysts address three fundamental questions concerning ethical practice: What is the right thing to do? What is worth doing? and What does it mean to be a good behavior analyst? The authors define ethics; outline standards of professional practice and a code of conduct for behavior analysts; describe strategies and tactics for providing ethically sound client services; and recommend how behavior analysts can achieve, maintain, and extend their professional competence.

CHAPTER 31

Ethical and Professional Responsibilities of Applied Behavior Analysts

Thomas R. Freeman, Linda A. LeBlanc, and Jose A. Martinez-Diaz

LEARNING OBJECTIVES

- Define ethics and state why ethical practice and codes are important.
- Define and explain standards of professional practice relevant to behavior analysts.
- Explain how professional competence can be ensured.
- Identify and explain ethical issues in client services.
- Explain considerations important to advocating for clients.

Ethical dilemmas surface frequently in the practice of applied behavior analysis. Consider these situations:

- Lee resides in a private, rural, for-profit community-based home for persons with intellectual disabilities. He informs Dr. Han, the facility director, that he wants to move to an affordable apartment in a nearby town. Such a move would represent a loss of income to the agency, might generate additional transition costs (e.g., moving expenses, future on-site supervision), and has the potential to be dangerous, given the spike in the crime rate in the area of town that Lee is considering. How could Dr. Han respond ethically to Lee's request about moving without being biased by a conflict of interest?
- Raluca has begun her supervised fieldwork in preparation for becoming a Board Certified Behavior Analyst (BCBA®). Her supervisor, Anca, is already a BCBA. Recently, Anca has directed Raluca to do more advanced tasks (e.g., assessments, designing programs) that are beyond the material that she has encountered in her coursework. She has not been explicitly taught how to do the tasks that Anca is directing her to complete with clients. Further, given circumstances, Anca is not directly observing her work or providing specific feedback. Raluca is concerned that she will make errors that could impact the quality of her client's programming. What should Raluca do to ensure that she is learning new behavior analytic skills fluently and that her clients have optimal programming? What should Anca have done to mitigate this situation from developing in the first place?
- Ms. Dougherty is a second-year teacher who is also pursuing experience hours toward certification in behavior analysis. During an annual individualized education program (IEP) meeting, she perceives that a school district administrator, Mr. Doyle, is trying to "steer" the parents of a student with emotional disabilities into accepting a revised IEP. Mr. Doyle has removed the recommendation from the team for applied behavior analysis services and replaced it with his own suggestion for sensory-based calming techniques to be provided by a paraprofessional. Ms. Dougherty believes Mr. Doyle's suggestion may be motivated, at least in part, by his desire to contain costs for his tightly budgeted school district, where professional behavior analysis services would represent a costlier intervention for this student. As a second-year teacher, Ms. Dougherty is concerned that if she speaks up, she might upset her principal and possibly lose her job. If she remains silent, however, and does not inform the parents of the team's original recommendation, the student is unlikely to receive needed services. Since she is accruing supervised experience hours, and is, therefore, responsible for behaving in accordance with the Compliance Code of the BACB, how might Ms. Dougherty serve as an effective advocate for the student, while attempting to retain both her faculty position and her credibility with school administrators?

How should the behavior analyst respond in each of these situations? This chapter will help to address that question. We begin by defining ethics and discussing the importance of maintaining ethical and professional standards in the science and practice of applied behavior analysis. We discuss ethical issues within the context of client services (e.g., informed consent, conflict of interest), ethical implications of new technologies, and the need to create professional networks to support ethical behavior.

We thank Timothy E. Heron for his input, and acknowledge Matthew Normand for his contributions to a previous version of this chapter.

WHAT IS ETHICS AND WHY IS IT IMPORTANT?

Ethics Defined

From a behavior analytic viewpoint, **ethics** is defined by specific rules of conduct. These rules can either be broadly stated, as in Koocher and Keith-Speigel's (1998) guiding principles—(1) do no harm; (2) respect autonomy; (3) benefit others; (4) be just; (5) be truthful; (6) accord dignity; (7) treat others with care and compassion; (8) pursue excellence; and (9) accept accountability—or more specifically delineated, as in a professional code of conduct. For applied behavior analysts, The Professional and Ethical **Compliance Code** for Behavior Analysts—hereafter referred to as *The Compliance Code* (see Figure 31.1)—includes 10 sections relevant to professional and ethical behavior across an array of areas.[1]

Figure 31.1 The Professional and Ethical Compliance Code of the Behavior Analyst Certification Board.

1.0 Responsible Conduct of Behavior Analysts
- 1.01 Reliance on Scientific Knowledge
- 1.02 Boundaries of Competence
- 1.03 Maintaining Competence through Professional Development
- 1.04 Integrity
- 1.05 Professional and Scientific Relationships
- 1.06 Multiple Relationships and Conflicts of Interest
- 1.07 Exploitative Relationships

2.0 Behavior Analysts' Responsibility to Clients
- 2.01 Accepting Clients
- 2.02 Responsibility
- 2.03 Consultation
- 2.04 Third-Party Involvement in Services
- 2.05 Rights and Prerogatives of Clients
- 2.06 Maintaining Confidentiality
- 2.07 Maintaining Records
- 2.08 Disclosures
- 2.09 Treatment/Intervention Efficacy
- 2.10 Documenting Professional Work and Research
- 2.11 Records and Data
- 2.12 Contracts, Fees, and Financial Arrangements
- 2.13 Accuracy in Billing Reports
- 2.14 Referrals and Fees
- 2.15 Interrupting or Discontinuing Services

3.0 Assessing Behavior
- 3.01 Behavior-Analytic Assessment
- 3.02 Medical Consultation
- 3.03 Behavior-Analytic Assessment Consent
- 3.04 Explaining Assessment Results
- 3.05 Consent-Client Records

4.0 Behavior Analysts and the Behavior-Change Program
- 4.01 Conceptual Consistency
- 4.02 Involving Clients in Planning and Consent
- 4.03 Individualized Behavior-Change Programs
- 4.04 Approving Behavior-Change Programs
- 4.05 Describing Behavior-Change Program Objectives
- 4.06 Describing Conditions for Behavior-Change Program Success
- 4.07 Environmental Conditions that Interfere with Implementation
- 4.08 Considerations Regarding Punishment Procedures
- 4.09 Least Restrictive Procedures
- 4.10 Avoiding Harmful Reinforcers
- 4.11 Discontinuing Behavior-Change Programs and Behavior-Analytic Services

(continued)

Figure 31.1 (continued)

5.0 Behavior Analysts as Supervisors
- 5.01 Supervisory Competence
- 5.02 Supervisory Volume
- 5.03 Supervisory Delegation
- 5.04 Designing Effective Supervision and Training
- 5.05 Communication of Supervision Conditions
- 5.06 Providing Feedback to Supervisees
- 5.07 Evaluating the Effects of Supervision

6.0 Behavior Analysts' Ethical Responsibility to the Profession of Behavior Analysts
- 6.01 Affirming Principles
- 6.02 Disseminating Behavior Analysis

7.0 Behavior Analysts' Ethical Responsibility to Colleagues
- 7.01 Promoting an Ethical Culture
- 7.02 Ethical Violations by Others and Risk of Harm

8.0 Public Statements
- 8.01 Avoiding False or Deceptive Statements
- 8.02 Intellectual Property
- 8.03 Statements by Others
- 8.04 Media Presentations and Media-Based Services
- 8.05 Testimonials and Advertising
- 8.06 In-Person Solicitation

9.0 Behavior Analysts and Research
- 9.01 Conforming with Laws and Regulations
- 9.02 Characteristics of Responsible Research
- 9.03 Informed Consent
- 9.04 Using Confidential Information for Didactic or Instructive Purposes
- 9.05 Debriefing
- 9.06 Grant and Journal Reviews
- 9.07 Plagiarism
- 9.08 Acknowledging Contributions
- 9.09 Accuracy and Use of Data

10.0 Behavior Analysts' Ethical Responsibility to the BACB
- 10.01 Truthful and Accurate Information Provided to the BACB
- 10.02 Timely Responding, Reporting, and Updating of Information Provided to the BACB
- 10.03 Confidentiality and BACB Intellectual Property
- 10.04 Examination Honesty and Irregularities
- 10.05 Compliance with BACB Supervision and Coursework Standards
- 10.06 Being Familiar with This Code
- 10.07 Discouraging Misrepresentation by Non-Certified Individuals

Rules of ethics describe the personal behavior of individuals, not vague inherent traits or characteristics of individuals or groups. One should not refer to a person or an organization as being either ethical or unethical; these terms describe the behavior of an individual.

The two main approaches to ethics are termed the *deontological* (i.e., acts-oriented) approach and the *utilitarian* or *teleological* (i.e., results-oriented) approach. The deontological approach evaluates the ethicality of behavior based solely on the nature of the act itself regardless of consequences (O'Donohue & Ferguson, 2011; White, 1993). The utilitarian approach judges the rightness or wrongness of a behavior based on its consequences or results.

Both approaches are important for behavior analysts. The deontological approach creates a firm and necessary set of rules

to guide ethical behavior. For example, *The Compliance Code* states that behavior analysts must not accept gifts. However, the utilitarian approach provides needed flexibility when implementing those rules of conduct. In relation to the prohibition regarding acceptance of gifts, it is not difficult to envision a situation where an adult with a developmental disability excitedly presents her behavior analyst with a picture she has drawn, or a young client, shyly gives the behavior analyst a clay statue he made at school as a holiday gift. The behavior analyst may unnecessarily damage the clinical relationship by rejecting such gifts on the grounds of maintaining ethical purity. However, as Lattal and Clark (2007) stated, "Rules alone are neither specific enough nor comprehensive enough to cover all aspects of a difficult ethical decision" (p. 6). The utilitarian approach thus provides a principled way to analyze and resolve ethical questions that are not answered with a simple "yes" or "no." One might consider the deontological approach to be analogous with "the letter of the law," while the utilitarian approach attempts to address "the spirit of the law."

From a philosophical perspective, Skinner (1978, 2005) tended toward the utilitarian view, repeatedly suggesting that ethics govern behavior, and are therefore functionally related to the survival of the individual, group, culture, and species (de Melo, de Castro & de Rose, 2015). When various interests across groups or cultures are not in full agreement, or when different ethical rules conflict, complexities arise. For instance, a behavior that might serve the long-term survival interests of a group may nonetheless work to the short-term detriment of an individual. When a behavior analyst is confronted with a complex ethical question that presents competing, contradictory, or potentially harmful contingencies, effective navigation toward a fair outcome requires a well-calibrated ethical compass and reliable anchor. The behavior analyst must carefully consider who is likely to benefit, who might be potentially harmed, and what guidance is provided by the professional and ethical code of conduct.

Behavior analysts' decisions and actions must always be ethically grounded in the answers to three fundamental questions: What is the right thing to do? What is worth doing? and What does it mean to be a good behavior analyst? (Reich, 1988; Smith, 1987, 1993). The answers to these questions guide each behavior analyst's personal and professional practices based on an overriding principle: to help others learn how to improve their physical, emotional, social, and overall personal condition. As Corey, Corey, and Callanan (1993) stated, "The basic purpose of practicing ethically is to further the welfare of the client" (p. 4).

What Is the Right Thing to Do?

The answer to this question is influenced by many factors, including personal learning histories of right and wrong, the context within which applied behavior analysis is practiced, and the applicable ethical and legal rules of conduct. One makes a determination based on established principles, validated methods, and contact with effective decisions made by other applied behavior analysts and professionals in related fields. The goal is to ensure the welfare of those entrusted to one's care, the integrity and robust growth of the profession, and ultimately the survival of the culture (Skinner, 1953, 2002).

Ethics is ultimately related to cultural practices. Ethical codes may vary between cultures or within a single culture over time. What is acceptable in one culture may not be acceptable in another (e.g., recreational consumption of alcohol). What is acceptable at one point in time may be completely unacceptable 20 years later.

Personal Histories. Each person represents a distinctly different learning history. Thus, each professional should strive to obtain initial and recurrent training and experiences to offset any bias or predisposition that may carry over from our personal or cultural background. For example, a behavior analyst who grew up watching her brother engage in self-injurious behavior (SIB) is now providing services to a child who exhibits similar SIB. She may recall her parents' use of various procedures to help her brother (e.g., use of punishment, exclusion from family outings, psychotropic medication). The potential influence of such personal experience can be reduced effectively, in favor of an unbiased exploration of clinically appropriate treatment alternatives, if the analyst has had proper training in a range of assessment and treatment options.

A practitioner's cultural or religious background may also influence decisions about the right course of action. An individual reared in a family culture based on a "spare the rod and spoil the child" philosophy may react differently to a child's tantrum than if that same individual was raised by parents who took a "he'll grow out of it" approach. With proper training in behavior analysis, his response will likely be less culturally biased and more reflective of a scientific evidence-based approach.

Finally, a person's professional training and clinical experience can influence decision making through such subtle biases as the practitioner's comfort with the use of familiar treatment options. This may lead to decisions that extend beyond proper discriminative control (i.e., the same interventions are used in cases where those treatment options are not best practice). Thus, the practitioner must remain vigilant, aware that personal history, cultural background, and even clinical experience may exert undue influence on the proper delivery of professional service. To counteract these influences, the practitioner should regularly discuss professional issues with colleagues, obtain training, seek supervision or consultation when needed, read recent research literature, and review case studies relating to one's area of professional focus (Bailey & Burch, 2016). If undue influence still presents a problem, the ethical practitioner, after consultation, should consider transferring the case to another provider.

The Context of Practice. Applied behavior analysts work in schools, homes, community settings, job sites, centers, and other environments. Rules within these environments cover myriad response classes (e.g., attendance, use of sick time, use of specific positive or aversive intervention strategies). Included in the rules are policy statements that are designed to help practitioners discriminate between legal issues and ethical issues. For example, some practices may be legal, but unethical. Breaking a professional confidence, accepting valued family "heirlooms" in lieu of payment for services, or engaging in consensual sexual relations with a recently discharged client over age 18 serve as examples of actions that are legal, but unethical. Other actions

are both illegal and unethical. For instance, misrepresenting personal skills or promised services; overbilling for services; stealing the belongings of a client while rendering services; abusing a client physically, emotionally, sexually, or socially; and engaging in consensual sexual interaction with persons under the legal age of consent are all examples of actions that are both illegal and unethical (Greenspan & Negron, 1994). Behavior analysts who can discriminate between these legal and ethical distinctions are more likely to provide effective services, maintain a focused sensitivity toward clients, and not run afoul of the law or professional standards of conduct.

Ethical Codes of Behavior. Professional organizations and associations generate or adopt **ethical codes of behavior**, to provide clear guidelines for their members to follow when deciding a course of proper action in conducting their professional duties. These codes also typically outline the rules and standards by which graduated sanctions may be imposed on members who deviate from the professional code (e.g., reprimand, censure, expulsion from the organization). The Association of Professional Behavior Analysts has adopted the BACB's *Professional and Ethical Compliance Code for Behavior Analysts* (Behavior Analyst Certification Board [BACB], 2014). More discussion of this point is found later in the chapter.

What Is Worth Doing?

Questions related to what is worth doing directly address the goals of practice (i.e., what we hope to accomplish) and the impact of the strategies used to pursue those goals. The answers to these questions are typically based on an evaluation of the existing exigencies, the need for intervention, and an analysis of social validity and cost–benefit ratios that relate to any proposed intervention.

Existing Exigencies. An adult who engages in severe self-injurious behavior, a child with significant feeding problems, or a student who engages in highly disruptive behavior in class each presents a distinct problem-solving challenge to the behavior analyst. Behaviors that are self-injurious or that harm others are unquestionably worthy of swift assessment and intervention. Such dangerous behaviors must ethically be considered as targets for change before other less dangerous or disruptive behaviors are addressed. However, the need to expedite treatment is not an invitation to adopt a form of *situational ethics*, in which the promise of a quick result in the short term compromises one's consideration of the long-term effects of the intervention. For example, a positive punishment procedure may produce an immediate reduction in behavior, but had a careful assessment of the function of the behavior been completed, a less intrusive intervention could have been used. Behavior analysts must, therefore, strike a balance between immediate health and safety concerns, any likely short-term and long-term effects of an intervention, and insuring that any behavioral procedures comply with *The Compliance Code*.

Social Validity. Questions of social validity ask (a) whether the goals of the planned intervention are acceptable; (b) whether procedures are acceptable and aligned with best treatment practices (Peters & Heron, 1993); and (c) whether the obtained results show meaningful, significant, and sustainable change (Wolf, 1978). Most people would agree that teaching a school-age child to read is a desirable goal. Using a measurably effective instructional method such as direct instruction is procedurally sound. Demonstrating marked improvement in reading skill is a socially significant outcome. In every sense, teaching a child or adult to read meets the ethical test for social validity. The new reading skill has a positive effect on the life of the individual.

However, we cannot make the same claim for every skill in every situation. Consider the following real-life examples: A 50-year-old woman with mobility, sight, and hearing problems is being taught to identify various street signs; an elderly patient with Alzheimer's disease is being taught to stack colored rings; a first-grade student with autism receives daily discrete trial instruction on discriminating various sports cars. In each of these cases, the present and future "worth" of the treatment with respect to resources, goals, and outcomes is questionable at best. From an ethical standpoint, the behavior analyst must carefully consider whether a particular goal should be accomplished, and whether the achievement of that goal would be sustainable in the natural environment.

Cost–Benefit Ratio. A cost–benefit analysis is contextual. The behavior analyst must weigh the likely difference between the cost of a treatment or intervention (i.e., resources involved in planning, implementing, monitoring, and evaluating) and the projected benefits (i.e., likely future potential gains by the person). The benefit to the individual should be significant enough to justify the short- and long-term cost of providing the service. For example, we would likely find it ethically difficult to support a plan to transport a student with learning disabilities to an expensive out-of-state private school when comparable services are available as part of the local neighborhood public school system, especially if the move to a different school would not clearly result in significantly improved learning and social behavior outcomes. In another case, an 11th-grade student with developmental disabilities might not be best served by being placed in an academic, college-oriented program, when a more functional individualized curriculum would be far more likely to increase this student's self-sufficiency, independence, and long-term prospects for employment. We would be acting unethically to assign this student to an environment where an overly ambitious academic program is likely to result in escape-related reflexive establishing operations, thereby evoking escape/avoidance behavior—especially when a more functional educational model would effectively help this student meet his individualized short-term and long-term goals. Sprague and Horner (1991) suggested that committees are best suited to address the thorny issues that often arise in a cost–benefit analysis. A committee-based decision-making process can effectively account for a wide range of perspectives that are placed into a hierarchy. This can help ensure that the opinions of those persons with the highest stake in the outcome are given the greatest consideration.

What Does It Mean to Be a Good Behavior Analyst?

Being a good behavior analyst requires more than following professional codes of conduct. While adherence to *The Compliance Code* is necessary, it is not sufficient. Moreover, keeping the client's welfare at the forefront of one's decision-making process is vitally important, but not, in and of itself, sufficient. A good practitioner must be self-regulating (Rosenberg & Schwartz, 2018). This means that the ethical practitioner seeks ways to calibrate decisions over time, making sure to take into consideration a well-informed integration of values, contingencies, rights, and responsibilities (Smith, 1993). In addition, good behavior analysts maintain contact with the evolving professional literature, seek advanced training to improve their skills and practices using a systematic method (Rosenberg & Schwartz, 2018).

Why Is Ethics Important?

Behavior analytic practitioners abide by ethical principles to (a) produce socially significant behavior change that produces contextually relevant and meaningful outcomes for the persons entrusted to their care (Hawkins, 1984), (b) reduce or eliminate specious or questionable interventions and the potential for harm (e.g., poor treatments, injuries to self or others), and (c) conform to the ethical standards of learned societies and professional organizations (https://bacb.com/ethics-code/). Skinner (1953) was concerned with the development of ethics, particularly in relation to those in society who require care and protection. His ethical perspective focused in part on the influence of **countercontrol**, which he defined as occurring when someone responds to coercion or an aversive form of external control by showing "an emotional reaction of anger or frustration including operant behavior which injures or is otherwise aversive to the controller" (p. 321).

Countercontrol is often characterized by the individual "moving out of range, attacking, or passively resisting" (Delprato, 2002, p.192). Skinner stated that countercontrol is naturally evoked in individuals who have the capacity to rebel against exploitative or harsh controlling conditions, but that this capacity is limited or absent in certain groups, such as the elderly, children, individuals with an intellectual disability or a chronic mental health condition, or individuals who are incarcerated. (Skinner, 1976). O'Donohue and Ferguson (2011) extend this conceptualization, concluding that, "in those circumstances where individuals are disadvantaged, such that they have no effective mechanisms of countercontrol, external agencies need to step in to offset this imbalance" (p. 495), which occurs by developing and enforcing ethical codes of conduct.

Ethical codes not only provide guidance on what to do but also help practitioners avoid engaging in "situational ethics" whereby one acts (or does not act) based on expediency, pressure, or misplaced priorities (Maxwell, 2003). By following professional ethical codes, the practitioner increases the likelihood that he or she will deliver appropriate and effective services, which will produce a cumulative social benefit. Over time, these codified rules of conduct are refined through a selection-based process that serves to promote cultural survival. Skinner directly addresses this broader perspective on the importance and function of ethical rules in his discussions on group control and cultural selection.

> Without help a person acquires very little moral or ethical behavior under either natural or social contingencies. The group supplies supporting contingencies when it describes its practices in codes or rules, which tell the individual how to behave, and when it enforces those rules with supplementary contingencies . . . This is all part of the social environment called culture, and the main effect, as we have seen, is to bring the individual under the control of remoter consequences of his behavior. The effect has had survival value in the process of cultural evolution, since practices evolve because those who practice them are as a result better off. (Skinner, 2002, p. 173)

Practitioners should recognize that behaving ethically may not result in direct and immediate reinforcement. Ethical action may in fact disadvantage the practitioner in the short term. For example, a behavior analyst may decline a highly promising business opportunity because it represents a potential conflict of interest, or even the appearance of a conflict. In rejecting this financial opportunity, the behavior analyst exhibits behavior that has come under the control of a set of rules whereby he sets aside immediate gain, choosing instead to abide by ethical standards that, in the long run, elevate the entire community. According to Skinner, such actions promote the survival of the group and, by extension, its cultural practices. Moreover, the benefit to the group redounds back upon the individual behavior analyst in the long run, since "reasonable control over the selfish behavior of the individual is matched by the advantages which he gains as a member of a group which controls the same selfish behavior in others" (Skinner, 1953, p. 327).

STANDARDS OF PROFESSIONAL PRACTICE FOR APPLIED BEHAVIOR ANALYSTS

What Are Professional Standards?

Professional standards are written guidelines or rules that provide direction for conducting the practices associated with an organized profession. Professional societies and certification or licensing boards develop, refine, and revise the standards that govern their profession to provide members with parameters for appropriate behavior in a dynamic and changing environment. In practice, organizations initially form task forces to develop standards that are reviewed and approved by their respective boards of directors and members. In addition to prescribing rules of conduct, most professional organizations exact sanctions on members who fail to follow the rules. Major code violations may result in expulsion from the organization, or revocation of certification or licensure.

Five complementary and interrelated documents describe standards of professional conduct and ethical practice for applied behavior analysts. Two of these documents specify codes of conduct (*Professional and Ethical Compliance Codes for Behavior Analysts, Ethical Principles of Psychologists and Code of*

Figure 31.2 The right to effective behavioral treatment.

The Right to Effective Behavioral Treatment

1. An individual has a right to a therapeutic environment.
2. An individual has a right to services whose overriding goal is personal welfare.
3. An individual has a right to treatment by a competent behavior analyst.
4. An individual has a right to programs that teach functional skills.
5. An individual has a right to behavioral assessment and ongoing evaluation.
6. An individual has a right to the most effective treatment procedures available.

Adapted from "The Right to Effective Behavioral Treatment," by the Association for Behavior Analysis International, 1989. Retrieved April 15, 2017, from https://www.abainternational.org/about-us/policies-and-positions/. Copyright by the Association for Behavior Analysis International. Adapted with permission.

Conduct). Three additional documents are position statements (i.e., *The Right to Effective Behavioral Treatment, Statement on Restraint and Seclusion, The Right to Effective Education*). These documents were published on the following dates.

- *Professional and Ethical Compliance Code for Behavior Analysts* (Behavior Analyst Certification Board [BACB], 2014).
- *Ethical Principles of Psychologists and Code of Conduct* (American Psychological Association, 2010).
- *The Right to Effective Behavioral Treatment* (Association for Behavior Analysis International, 1989).
- *Statement on Restraint and Seclusion* (Association for Behavior Analysis International, 2010).
- *The Right to Effective Education* (Association for Behavior Analysis International, 1990).

Professional and Ethical Compliance Code for Behavior Analysts

According to the preamble in the *Professional and Ethical Compliance Code for Behavior Analysts*, this document "consolidates, updates, and replaces the former BACB's Professional Disciplinary and Ethical Standards and Guidelines for Responsible Conduct for Behavior Analysts" (2014, p. 1). *The Compliance Code* contains 71 elements organized into 10 different sections, "relevant to professional and ethical behavior of behavior analysts, along with a glossary of terms." With an effective date of January 1, 2016, "all BACB applicants, certificants, and registrants will be required to adhere to the Code" (2014, p. 1). (See Figure 31.1.)

The Compliance Code has undergone several revisions over the past decade, changing from unenforceable guidelines, references, and resources to an enforceable document with disciplinary standards. **Disciplinary standards** are statements that affect intellectual property, malpractice, and an expanded compliance system for providing corrective, disciplinary, and revocation actions, depending on the circumstances (BACB, 2014; Carr, Ratcliff, Nosik, & Johnston, in press). The entire *Compliance Code* not only serves as a reference for clients and employers in determining whether a behavior analyst's behavior has been conducted in an appropriate, professional, and ethical manner (Carr et al., in press), but also provides a procedure for the supervised practice of aspiring behavior analysts, handling complaints, and enforcing the code of ethics.

Figure 31.3 Principal areas of endorsement from the ABAI Statement on Restraint and Seclusion.

1. The Welfare of the Individual Served is the Highest Priority
2. Individuals have a Right to Choose
3. The Principle of Least Restrictiveness

Adapted from "Statement on Restraint and Seclusion, 2010," by the Association for Behavior Analysis International. Retrieved October 30, 2017, from https://www.abainternational.org/about-us/policies-and-positions/. Copyright by the Association for Behavior Analysis International. Adapted with permission.

The Right to Effective Behavioral Treatment

The Association for Behavior Analysis International has published two position papers describing client's rights. In 1986, the Association established a task force to examine the rights of persons receiving behavioral treatment and how behavior analysts can ensure that clients are served appropriately. After 2 years of study, the task force outlined six basic client rights as the basis for directing the ethical and appropriate application of behavioral treatment (see Figure 31.2) (Van Houten et al., 1988).

Statement on Restraint and Seclusion

Figure 31.3 shows a portion of the Association for Behavior Analysis International's statement on restraint and seclusion. Overall, this statement underscores ABAI's position that the principles of least restrictiveness, the individual's right to choose, and the overall welfare of the individual are the highest priority.

ENSURING PROFESSIONAL COMPETENCE

Professional competence in behavior analysis is achieved through academic training that involves formal coursework, supervised practica, and mentored professional experience. Even though very few universities house departments of behavior analysis, many excellent behavior analysts have been trained through university-based master's and doctoral-level programs housed in psychology, education, and other human service departments.[2]

Behavior analysis accreditation and credentialing bodies have specified minimum curriculum and supervised experience requirements for behavior analysts. Both the Behavior Analyst Certification Board (BACB) and the Association for Behavior Analysis International Accreditation Board have set minimum standards for what constitutes training in behavior analysis. The BACB credentials individual practitioners, while ABAI Accreditation Board accredits university training programs and verifies that course sequences meet the instructional requirements for certification as set by the BACB. Practitioners must not only meet the certification criteria but also pass a certification examination. The BACB has conducted an extensive occupational analysis and developed the Behavior Analyst Task List, now in its fourth edition—Behavior Analyst Certification Board (BACB), 2012. The Task List specifies the minimum content that all behavior analysts should master (BACB, 2012; Shook, Johnston, & Mellichamp, 2004; cf. Martinez-Diaz, 2003).[3] The fifth edition of the Task List (BACB, 2017) becomes effective January 1, 2022.

Obtaining Certification and Licensure

Potential consumers must be able to identify practicing behavior analysts who have demonstrated at least a minimum standard of training and competence (Moore & Shook, 2001; Shook & Favell, 1996; Shook & Neisworth, 2005). A professional credential publicly identifies those who have met certain requirements in a field (Shook & Favell, 1996). Those seeking professional services often use a profession's credential as a tool in selecting a provider, and employers may use the credential as an efficient means of hiring professionals (Carr et al., in press). This information, combined with the disciplinary action that can occur if a professional acts outside the scope of responsible professional conduct, is designed to enhance public protection. In the past, most behavior analysts in private practice were licensed in psychology, education, or clinical social work. The public had no way to determine whether the licensed professional had specific training in applied behavior analysis (Martinez-Diaz, 2003).

Two distinct categories of professional credentials are important in the field of applied behavior analysis: professional certification and licensure.

Professional Certification

Professional certification is a voluntary credential issued by a private organization that is authorized by members of the profession, not the government. The professional certifying body for applied behavior analysis is the BACB, whose board of directors is composed of behavior analysts and a consumer representative. This international credential crosses state and national borders, offering greater mobility to members of the profession whose credentials might otherwise be restricted to specific areas (Hall & Lunt, 2005). The eligibility requirements for certification are predicated upon educational degree and coursework requirements, supervised experiential training requirements, and passing a psychometrically sound examination (Johnston, Mellichamp, Shook & Carr, 2014). Like other professions, the BACB requires documented continuing educational experiences and professional development, with specific emphasis on the areas of ethics and supervision (Carr et al., in press).

State/National-issued Licenses

A license is a state, provincial, or national government document that is established by law and specifies the privileges and limits of professional practice (Green & Johnston, 2009). A license is a mandatory credential needed to practice, not a voluntary certificate of achievement. Standards for obtaining, maintaining, and legally revoking a license are developed by the designated regulatory board that may or may not include members of the profession. As laws are constantly changing within and across states or nationally, behavior analysts are responsible for remaining up to date with respect to any licensing regulations that govern their practice within that jurisdiction.

Practicing Within One's Areas of Competence

Behavior analysts must practice within their areas of competence, based on professional training, experience, and performance. For example, a behavior analyst with extensive experience working with typically developing adolescents should not begin working with children with autism-related language delays or adults with developmental disabilities without obtaining additional active supervision. The behavior analyst must accrue adequate supervised experience working with any new clinical population. Moreover, an individual with experience only in clinical practice should not begin to provide services in the area of organizational behavior management without additional training and supervision. LeBlanc, Heinicke, and Baker (2012) provide guidance on expanding areas of practice in an ethically responsible manner.

Even when practicing within one's area of competence, situations may arise that exceed one's training or experience. In these cases, the practitioner must either make a referral to another behavior analyst or consultant, or work in close consultation with another provider who has the necessary expertise. Gaps in one's professional training may be addressed by attending workshops, seminars, classes, and other continuing education activities. However, professional situations can arise that are outside a BCBA's experience (see Figure 31.4). At that point, behavior analysts should work with mentors, supervisors, or colleagues who can provide enhanced training, professional development, and guidance.

Developing Competence in Supervision and Mentorship

Recognizing the Need

Behavior analysts often provide supervision for (a) young professionals who are pursuing certification as a behavior analyst, (b) individuals holding paraprofessional-level credentials (e.g., Registered Behavior Technician®), (c) certified behavior analysts who require supervision of their ongoing practice, and (d) non–behavior analysts. Effective supervision of others is critical to the quality of ongoing behavioral services. It contributes to the professional development of the supervisee, the continued growth of the supervisor, and the overall development of our field and its practice, and is recognized in the professional literature (Sellers, Valentino, & LeBlanc, 2016). The supervisor who also fully embraces the role of mentoring is positioned to serve as a life-long counselor through the many challenging professional situations

Figure 31.4 Seeking guidance to handle an unexpected ethical situation.

Presenting Concern	Ethical Problem-solving Strategies
Background: Rhonda recently became employed as a senior behavior analyst for the state. She chairs the local program review committee in District 20, and monitors behavioral services delivered under state authorizations. Voting members of the committee must be *Board Certified Behavior Analysts* (BCBAs). Stanley, a long-time provider in District 20, who was grandfathered under the old state certification rules, always signed his programs in the past as "Certified Behavior Analyst." However, in a recent presentation, Stanley signed his program (which was countersigned by the parents), "Stanley Smith, BCBA." After the program review committee meeting, Rhonda approached Stanley and asked, "Stan, I see that you have signed off as a BCBA. Did you pass the certification exam recently?" Stan answered, "I sent in my certification. You should have it in my file." Rhonda searched and checked the committee files, and other appropriate files, but did not find Stanley's certificate. Rhonda then asked Stanley to provide a copy of his certification, on the chance that it was "lost in the system." He did not comply. Then, she checked the board registry and he was not listed. She asked him to come into the office for a meeting to clarify the situation. He declined. Stanley then filed an age discrimination complaint with the state Equal Employment Opportunity Commission office. What would be Rhonda's best course of action to address this situation?	*Analysis*: Stanley does not appear to be a BCBA, and therefore should not be signing documents as if he were. He did not provide documentation of his certification, he is not listed on the public registry, and he declined to meet with Rhonda to explain the discrepancy between his statement ("I sent in my certification") and the fact that he is not listed in the registry. He is likely in violation of at least *The Compliance Code* 1.04 Integrity, 7.02 Ethical Violations by others and risk of harm, 8.01 Avoiding false or deceptive statements, and possibly 1.02 Boundaries of competence. *Action*: Before taking further action that might escalate the matter, Rhonda would be wise to do two things: (1) She should attempt to verify that Stan indeed does not have a certificate of any kind that could have been legitimately issued in the past under the grandfather clause (basically ruling out that Stan's certificate was lost in the mail or the file somehow); and (2) she should seek guidance from her supervisor in the state service organization to ensure that she follows the proper regulations in dealing with this potential major violation. Under *The Compliance Code* 10.07 Discouraging misrepresentation by noncertified individuals, Rhonda has an ethical responsibility to report Stanley to the BACB immediately. Moreover, Rhonda may have to initiate an investigation by the state of Stanley's billing practices, to be sure that he has been billing appropriately for the level of services he is providing. This may require her to report Stanley directly to the fraud division of a funder. In terms of the discrimination complaint, she would be wise to save and file all documentation of her interactions with Stanley, beginning with the date of her first inquiry. Alerting her supervisor of the disposition of the complaint through *carbon* copies of e-mails or letters would help to ensure total transparency.

that may arise throughout a career. Serving in the role of supervisor requires training, and *The Compliance Code*'s changing standards underscore this need (LeBlanc & Luiselli, 2016).

In response to the need to provide competent training, the BACB convened a workgroup on supervision that recommended several requirements for all certificants providing supervision for those already credentialed, or those pursuing credentials. The requirements were implemented in 2015 and included (a) specific training in supervisory practices prior to supervising others, (b) continuing education in the area of supervision, and (c) a registry that publicly links the BCBA supervisor to his or her supervisee(s).

A special section of the journal *Behavior Analysis in Practice* includes several papers that outline practice recommendations for individual supervision, group supervision, and addressing problems that can arise during the supervisory relationship (LeBlanc & Luiselli, 2016; Sellers, LeBlanc, & Valentino, 2016; Sellers, Valentino, & LeBlanc, 2016; Turner, Fischer, & Luiselli, 2016; Valentino, Sellers, & LeBlanc, 2016).

Developing the Supervisor's Skill Set

The supervisor bears broad responsibility in the supervisory relationship, including the obligation to shape core behavior analytic, conceptual, experimental, and applied repertoires (Sellers, Valentino, & LeBlanc, 2016; Turner et al., 2016). These professional skill sets also include social effectiveness training, public speaking to multiple audiences, time and stress management skills, problem-solving skills, and ethical decision making, among many others (Bailey & Burch, 2010).

Developing these repertoires through in-service training, mentoring experiences, or field-based practica, allows the behavior analyst to be more effective in influencing those who might adopt, implement, or disseminate effective behavioral programming. When these repertoires are actively modeled and explicitly taught to the supervisor-in-training, the likelihood of future professional success of her clients is increased. When these repertoires are compromised, the supervisee does not learn important skills, he may leave the professional altogether, the supervisory relationship may become strained or damaged, and the overall perception of behavior analysis may be hurt (Sellers, Valentino, & LeBlanc, 2016). Sellers, LeBlanc, and Valentino (2016) offer specific suggestions for detecting and remediating these issues throughout the supervisory period (see Figure 31.5).

Section 5.0 of *The Compliance Code* addresses the specific ethical guidelines for supervision by behavior analysts. This section establishes that the "supervisor must take full responsibility for all facets of this undertaking." Sellers, Alai-Rosales, and MacDonald (2016) provide a rationale for each sub-code on supervision with examples of the concerns that might arise if the behavior analyst does not take appropriate action to comply

Figure 31.5 Identifying and managing performance issues in a supervisory context.

Presenting Concern	Ethical Problem-solving Strategies
Background: Jasmina is a BCBA-D who supervises an aspiring behavior analyst, Barry, in his accrual of supervised fieldwork experience hours. Barry has been learning many new concepts and principles along with specific applied procedures to use with clients. Barry implements procedures accurately during therapy sessions, but Jasmina has noticed that Barry is often late to therapy and supervision sessions, makes many errors in data collection, and does not reliably turn in written work products on time. When he does turn in the work, it often has typographical errors or incomplete sentences. This week Jasmina realized that skill acquisition data on a client has not been graphed in several weeks. A casual review of the data sheets indicates that progress stalled over 3 weeks ago, and program changes are well overdue.	*Analysis*: Jasmina realizes that her supervisee has problems in attending, time management and organization skills, and this is affecting his ability to provide appropriate, data-based clinical services. They are failing in their basic ethical responsibilities to clients. While she may catch many of his errors by providing close supervision, she expects that Barry will continue to struggle with these issues once he is no longer closely supervised and is practicing independently. Jasmina identifies her ethical responsibility to her supervisee, the field, and the current and future clients of this supervisee. She must speak with Barry and create a professional development plan to remedy these concerns and help him establish the needed skills. As suggested by Sellers, LeBlanc and Valentino (2016), Jasmina should first reflect on whether she has clearly specified her performance expectations to Barry. If she has not yet given feedback about his late arrivals or poor written work, she should use basic instructions, modeling, rehearsal and feedback to explicitly address his specific problematic repertoires. If she has already provided training and feedback, and his behavior did not change, she may conclude that a more pervasive deficit in time management skills may be having an adverse impact on several of Barry's repertoires. *Action*: Jasmina should talk to him about this in clear and explicit terms, state her desire to help him succeed, and express her concern for the long-term interests of future clients if his problems are not resolved. She might ask whether he has had similar problems in the past. She might also ask to see his time management and scheduling system. If she determines that he has a more global problem with organization, which has a broad impact on his performance (e.g., no system for planning time and tracking tasks and deadlines), Sellers et al. provide recommendations to help develop a support plan. She could assign relevant readings (e.g., David Allen's, *Getting Things Done (2015)*); they could review readings together, and she could help Barry begin to employ new effective strategies. They could select an electronic schedule/notekeeping tool to establish a new system for Barry, with explicit reinforcement contingencies for success at each step in the development plan.

with *The Compliance Code*. For example, sub-code 5.02 refers to appropriate supervisory volume (i.e., a volume of supervisory activities, which is commensurate with the ability to provide high-quality, effective supervision leading to high-quality intervention services). Sellers, Alai-Rosales et al. provide a case example in which a well-intentioned BCBA gradually began to take on too many supervisees. Although her intent was to quickly build to capacity the BCBA-based services in her community, taking on more supervisees than she could effectively oversee resulted in a poor training experience for many of these new professionals. Whether a practitioner overcommits based only on financial pressures or due to a genuine intent to benefit the field, the risk for a poor outcome is equally high.

In another example, Sellers, Alai-Rosales et al. describe the importance of sub-code 5.05, which refers to the importance of communicating the conditions for supervision in writing at the outset of the supervisory relationship (i.e., in a supervision contract) and abiding by those communicated conditions. Conditions and expectations of the relationship must be clearly communicated when the relationship is established. If not, uncertainties arise, leading to disputes regarding required activities, performance criteria and evaluations, and whether or not a supervisor will sign off to verify that experience requirements have been met. If the supervisor and supervisee are not within the same organizational hierarchy (e.g., the supervisor does not directly oversee the clinical work of the supervisee), opportunities for misperception may be even more likely regarding valid activities, actions, and competencies required on the part of the supervisee. Clearly delineated conditions in a contract are critical in producing an effective working relationship. Many ethical dilemmas can arise in supervision, and the supervisor must fully accept the comprehensive responsibility that goes along with this role. The supervisor must proactively seek to ensure the best possible outcome for the supervisee, while ensuring delivery of the highest quality clinical services possible.

The ethics of developing a supervisor's skill set applies to the second scenario at the beginning of the chapter. Anca is not providing Raluca adequate supervision. She would have done far better had she communicated, in writing, the conditions and expectations for both herself and the supervisee at the outset of their supervisory relationship, based on the standards provided by the BACB—and followed those guidelines. Moreover, she should have assigned Raluca only to tasks for which she was adequately trained. At this point in the scenario, Raluca must inform Anca that she requires additional assistance, most likely in the form of hands-on supervision as required by the BACB's own guidelines. If Anca's supervision behavior does not immediately improve, Raluca will probably need to find a new, competent supervisor, and may need to report Anca to the BACB for providing inadequate supervision.

Maintaining and Expanding Professional Competence

Behavior analysts have an ethical responsibility to stay informed of advancements in the field. For example,

conceptual advances and technical innovations in areas such as functional analysis, motivating operations, and applications of the matching law have had profound implications for clinical and educational practice. Behavior analysts can maintain and expand their professional competence by earning continuing education credits, attending and participating in professional conferences, reading professional literature, and presenting cases to peer review and oversight committees.

Continuing Education Units

Behavior analysts can expand their professional competence and keep abreast of new developments by attending training events offering continuing education unit (CEU) credits. The BACB requires a minimum number of CEUs per recertification cycle to maintain certification. Types of continuing education include university coursework, CE issued by BACB-approved continuing education providers, instruction of either of the preceding types, CE issued by the BACB directly, scholarly activities such as publishing or reviewing articles for a journal, and taking and passing the BACB's certification exam again. Continuing education units demonstrate that the behavior analyst has added relevant awareness, knowledge, and/or skills to his or her repertoire. The BACB now requires CEUs in the areas of ethics and supervision during each renewal cycle, which is a testament to the growing importance of these two areas as critical to effective functioning of an applied behavior analyst.

Attending and Presenting at Conferences

Attending and participating in local, state, or national conferences enhance every behavior analyst's skills. The axiom "You never learn something so well as when you have to teach it" continues to have value. Hence, participating in a conference helps practitioners refine their skills.

Professional Reading

Self-study is a fundamental way to maintain currency in an ever-changing field. In addition to routinely reading *Behavior Analysis in Practice*, the *Journal of Applied Behavior Analysis* and *Perspectives on Behavior Science* (formerly published as *The Behavior Analyst*), all behavior analysts should study other behavioral publications specific to their areas of expertise and interest.

Oversight and Peer Review Opportunities

When behavior analysts approach a difficult problem, they apply skills and techniques within their repertoire to address the problem. For example, a behavior analyst may successfully fade out the use of a protective helmet with a child who has long engaged in frequent and severe face-slapping by using differential reinforcement of alternative behavior (DRA) to increase preferred-object holding and play activity by the child. Fading the safety-based helmet presents no ethical problems as long as the technique remains conceptually systematic, is tied to the published research literature, and is effective. However, even the most technically proficient and professionally careful practitioner is subject to contingencies that can lead to treatment drift or error. This is where oversight and peer review come into play.

Many states have laws requiring that oversight be provided under specific circumstances, which are often defined by the type and severity of behaviors being addressed and/or the restrictiveness or intrusiveness of the procedure(s) being proposed. However, the presence or absence of laws in a particular jurisdiction should not determine our dedication to a process of peer review and oversight. This type of professional monitoring protects (1) the consumers of behavior analytic services by ensuring the use of valid and appropriate procedures; (2) the individual behavior analyst, by providing both helpful feedback and administrative support and back-up; and (3) the field in general, in reducing the likelihood of lawsuits or scandal based on the rogue use of untested, inappropriate, and possibly dangerous treatment options.

When findings are presented before a group of colleagues, professionals, or other stakeholders outside the behavior analytic community, the practitioner must outline the use of procedures that comply with clinical and professional standards. Such presentations also provide the behavior analyst an opportunity to clarify behavioral outcomes, present graphic displays that are easy to interpret, and explain the rationale for the various educational or clinical choices in clear language, free of jargon.

Making and Substantiating Professional Claims

Sometimes overzealous or overly confident behavior analysts make preemptive claims that are not realistic. For example, asserting, "I am certain I can help your son" is almost certainly an unethical claim, especially prior to conducting a full functional assessment. A more ethical and appropriate statement may be, "I have had success working with other children who have profiles similar to your son's." A behavior analyst who is experienced in working with a particular population, skilled in the careful process of identifying functional relations, and well acquainted with the salient professional literature is less likely to make unsubstantiated or exceedingly ambitious claims.

This professional standard also relates to those situations where an individual makes a false or deceptive claim to having a certification, license, educational background, or type of training that the individual does not have. Making dishonest statements or claims relating to one's professional experience or credentials is always unethical and, as likely fraudulent conduct, may be illegal. This will be discussed more fully in the section below.

ETHICAL ISSUES IN CLIENT SERVICES

Although applied behavior analysis shares many ethical concerns with other disciplines, ethical considerations that are specific to behavior analysis do arise. Behavior analysts often address skills or problem behavior in contexts where privacy can be a significant concern (e.g., toileting, dressing) or where best-practice guidelines suggest the use of certain evidence-based, but potentially controversial, interventions (e.g., use of an aversive procedure to reduce a life-threatening behavior). Clinical decision making often poses ethical questions, which must be addressed prior to implementation (Herr, O'Sullivan, & Dinerstein, 1999; Iwata, 1988; Repp & Singh, 1990). Reading and referring to *The Compliance Code* and the Task List on a regular basis helps the practitioner maintain a strong ethical foundation.

Informed Consent

Any discussion of service delivery must begin with the issue of consent. A potential recipient of behavioral services, or a participant in a research study, must provide explicit written and signed permission before any assessment, treatment, or research participation may begin. This formal permission to begin any aspect of treatment—assessment, intervention, evaluation—is termed **informed consent**, and requires that full disclosure of all pertinent information be provided to the individual prior to a granting of permission. Figure 31.6 shows a basic example of an informed consent form in which such information is provided.

Three tests must be met before informed consent can be considered valid: (a) The person must demonstrate the *capacity to decide*, (b) the person's decision must be *voluntary*, and (c) the person must have adequate *knowledge* of all salient aspects of the treatment.

Capacity to Decide

To be considered capable of making an informed decision, a participant must have (a) an adequate mental process or faculty by which he or she acquires knowledge, (b) the ability to select and express his or her choices, and (c) the ability to engage in a rational process of decision making. Behavior analysts consider concepts such as "mental process or faculty" to be hypothetical constructs (i.e., mentalism), and there are no widely accepted evaluative tools for testing one's preintervention capacity. However, capacity is most likely to be questioned if an individual "has impaired or limited ability to reason, remember, make choices, see the consequences of actions, and plan for the future" (O'Sullivan, 1999, p. 13). A person is considered mentally incapacitated if a disability affects his or her ability to understand the consequences of his or her actions (Turnbull & Turnbull, 1998).

According to Hurley and O'Sullivan (1999), "Capacity to give informed consent is a fluid concept and varies with each individual and proposed procedure" (p. 39). A person may have the capacity to consent to a positive reinforcement program that poses little or no risk, but may not have the capacity to make decisions on more complex treatment alternatives such as the manipulation of motivating operations or use of punishment. Practitioners should not assume that a person who is capable of giving informed consent in one set of circumstances can provide truly informed consent in all cases.

Evaluation of capacity must be viewed within both a legal and a behavioral context. Courts have held that capacity, in relation to legal competence, requires "the subject to understand rationally the nature of the procedure, its risks, and other relevant information" (*Kaimowitz v. Michigan Department of Mental Health*, 1973, p. 150; Neef, Iwata, & Page, 1986). Thus, operationally, capacity is often determined in relation to the individual meeting the "knowledge" requirement (see below). The determination of capacity in persons with intellectual disabilities poses specific challenges, and the behavior analyst would benefit from consulting a legal expert whenever a question of capacity arises.

When a person is deemed incapacitated, informed consent may be obtained either through a surrogate or a guardian.

Surrogate Consent. *Surrogate consent* is a legal process by which another individual—the surrogate—is authorized to make decisions for a person deemed incompetent, based on the knowledge of what the incapacitated person would have wanted. Family members or close friends most often serve as surrogates.

Figure 31.6 Informed consent form.

(Name of agency/provider)

Informed Consent Form

CLIENT: ______________________________ DOB: __________

STATEMENT OF AUTHORITY TO CONSENT: I certify that I have the authority to legally consent to assessment, release of information, and all legal issues involving the above-named client. Upon request, I will provide (provider/agency) with proper legal documentation to support this claim. I further hereby agree that if my status as legal guardian should change, I will immediately inform (provider/agency) of this change in status and will further immediately inform (provider/agency) of the name, address, and phone number of the person or persons who have assumed guardianship of the above-named client.

TREATMENT CONSENT: I consent for behavioral treatment to be provided for the above-named client by (provider/agency), and its staff. I understand that the procedures used will consist of manipulating antecedents and consequences to produce improvements in behavior. At the beginning of treatment behavior may get worse in the environment where the treatment is provided (e.g., "extinction burst") or in other settings (e.g., "behavioral contrast"). As part of the behavioral treatment, physical prompting and manual guidance may be used. The actual treatment protocols that will be used have been explained to me.

I understand that I may revoke this consent at any time. However, I cannot revoke consent for action that has already been taken. A copy of this consent shall be as valid as the original.

PARENT/GUARDIAN: ______________________________ DATE: __________

WITNESS: ______________________________ DATE: __________

In most states, a surrogate's authority is limited. A surrogate may not authorize treatment when the client is actively refusing treatment (e.g., when an adult with developmental disabilities refuses to sit in a dentist's chair, a surrogate cannot consent to sedation); or for controversial medical procedures such as sterilization, abortion, or certain treatments for mental disorders (e.g., electroconvulsive therapy or psychotropic medication) (Hurley & O'Sullivan, 1999). The surrogate is required to consider specific information when making decisions. For a person who is incapacitated, the landmark case of *Superintendent of Belchertown State School v. Saikewicz* (1977) crystallized factors that surrogates must consider when making informed consent decisions for persons who cannot make their own decisions (Hurley & O'Sullivan, 1999). Figure 31.7 lists necessary information for the two primary areas of concern for surrogates: (a) decision making for a person who is incapacitated and whose wishes may be known or inferred and (b) decision making for a person whose wishes are unknown and probably unknowable (e.g., a person with a profound and lifelong developmental disability).

Guardian Consent. *A guardian is* an individual or organization named by order of the court to exercise any or all powers and rights over the person and/or estate of an individual (National Guardianship Association, 2016). Guardianship is a complex legal issue that varies by jurisdiction. Thus, we will address only two main points regarding guardianship.

First, guardianship may be limited in any way that the court deems appropriate. In most states, a full guardian is responsible for essentially every important decision in a person's life. The court, in protecting the rights of the individual, may determine that limited or temporary guardianship is more appropriate. For example, guardianship may apply only to financial or medical concerns, and may remain in effect only so long as a very specific concern is at issue (e.g., need for surgery). Guardians may not deny basic human rights, such as the right to freedom from abuse or the right to privacy.

Second, when treatment is deemed necessary for an individual who lacks capacity, but that individual refuses treatment, surrogate consent is insufficient. Only a guardian may provide oversight and consent in this situation. The greater the degree of guardianship, the less legal control a person has over his or her own life. Insofar as a major goal of behavior analysis is to increase the independence of those we serve, guardianship of any kind should be considered only when absolutely necessary, and at the most limited level that still resolves the issue.

In all guardianship cases, the court is the ultimate decision-making body and may take any action, including revoking guardianship or determining who should serve as the guardian (O'Sullivan, 1999).

Voluntary Decision

Consent is considered voluntary when it is given in the absence of coercion, duress, or any undue influence and when it is issued with the understanding that it can be withdrawn at any time. "Revocation of consent has the same effects as an initial refusal to consent" (Yell, 1998, p. 274).

Family members, doctors, support staff, or others may exert strong influence over a person's willingness to grant or

Figure 31.7 Factors surrogates must consider when making informed consent decisions for persons who are incapacitated.

<u>FOR A PERSON WHOSE WISHES MAY BE KNOWN OR INFERRED</u>

1. The person's current diagnosis and prognosis.
2. The person's expressed preference regarding the treatment issue.
3. The person's relevant religious or personal beliefs.
4. The person's behavior and attitude toward medical treatment.
5. The person's attitude toward a similar treatment for another individual.
6. The person's expressed concerns about the effects of his or her illness and treatment on family and friends.

<u>FOR A PERSON WHOSE WISHES ARE UNKNOWN AND PROBABLY UNKNOWABLE</u>

1. The effects of the treatment on the physical, emotional, and mental functions of the person.
2. The physical pain that the person would suffer from the treatment or from the withholding or withdrawal of the treatment.
3. The humiliation, loss of dignity, and dependency the person is suffering as a result of the condition or as a result of the treatment.
4. The effect the treatment would have on the life expectancy of the person.
5. The person's potential for recovery, with the treatment and without the treatment.
6. The risks, side effects, and benefits of the treatment.

Adapted from "Informed Consent for Health Care," by A. D. N. Hurley and J. L. O'Sullivan, 1999, in R. D. Dinerstein, S. S. Herr, and J. L. O'Sullivan (Eds.), *A Guide to Consent*, pp. 50–51. Washington DC, American Association on Intellectual and Developmental Disabilities.

deny consent (Hayes, Adams, & Rydeen, 1994). For example, individuals with developmental disabilities may be asked to make major decisions during interdisciplinary team meetings. The way questions are posed to the individual may suggest subtle or not-so-subtle pressure to provide consent (e.g., "This will really help you. You agree with all of us on this, right?"). Therefore, the practitioner should consider discussing issues requiring consent with the individual and an independent advocate in a smaller, private setting. The discussion should not be constrained by time limits, or indicate a need for a quick approval. The prospective service recipient should be given an opportunity to take time to think, discuss, and review all options with a trusted confidant. These steps help ensure that consent is voluntary.

Knowledge of the Treatment

A person considering behavioral services or participation in research must be provided information in clear, nontechnical language regarding: (a) all important aspects of the planned treatment, (b) all potential risks and benefits of the planned procedure or research protocols, (c) all potential evidence-based alternative treatments, and (d) the right to discontinue treatment at any time. Once information has been provided, the individual must demonstrate an understanding of that information, which clearly relates back to an assessment of that individual's capacity. The person must be able to describe the procedures in his or her own words, and correctly answer questions about the plan. For example, if a time-out procedure is suggested, the client should be able to describe how and when the time-out will be implemented. Instead of saying, "I'll get in trouble if I hit another person," the individual should say, "If I hit somebody, I have to go sit over there for 2 minutes, and then I can come back." Figure 31.8 outlines additional information that the behavior analyst should provide to the potential service recipient to ensure voluntary, informed consent (Cipani, Robinson, & Toro, 2003).

As part of providing information for consent, the behavior analyst should always do a **risk–benefit analysis**—that is, weighing the potential harm to the client (or others) versus the benefits likely to result from implementing those procedures—prior to conducting the assessment or intervention.[4] Bailey and Burch (2016) encourage conducting a risk–benefit analysis of four key areas: (1) the general risk factors for behavioral treatment, (2) the benefits of behavioral treatment, (3) the risk factors for each behavioral procedure, and (4) a reconciliation of the risks and benefits with key parties involved. This analysis should occur prior to initial or subsequent requests for informed consent, and repeatedly throughout the delivery of services. For example, a descriptive assessment has few risks, and can help determine factors that trigger or maintain behavior. However, it does not provide a clear demonstration of functional relations.

Alternatively, a full functional analysis, in which antecedents and consequences are systematically manipulated, can clearly demonstrate functional relations, but is also likely to evoke a short-term increase in the problem behavior, with limited potential for triggering a major behavioral episode, possible injury, or other iatrogenic (i.e., intervention-induced) effects.

A risk–benefit analysis must always be done when choosing one evidence-based treatment option over another. This type of analysis assists the practitioner in choosing the least restrictive option that is most likely to be effective.

Treatment Without Consent

Most states have policies authorizing a course of action when an individual cannot, or will not, provide informed consent for a necessary treatment. Typically, consent may be granted in the case

Figure 31.8 Information for clients to ensure informed consent.

The behavior analyst helps to ensure the validity and clarity of informed consent by completing the following tasks:

- Generates appropriate documents that are understandable in the person's primary language and that make clear all elements of the program, including who will deliver the services, their qualifications, under what conditions services will be rendered, the length of the proposed intervention, risks and benefits, and how the program will be evaluated.
- Ensures that all consent documents are discussed in a private setting, and, if necessary, with a family member, third-party advocate, or other witness present, as chosen by the service recipient, who may provide counsel and/or assistance.
- Schedules the consent meeting at a time and location that is convenient for the service recipient. This meeting should allow adequate time for all questions to be addressed, and for adjustments or alternatives to the proposed treatments to be considered.
- Documents in writing the responsibilities of the behavior analyst and the service recipient, including: How confidentiality will be handled (i.e., who has access to records); provisions for a renewed consent if, for example, new data emerge that have the potential to change the proposed intervention; and any issues associated with remuneration, communication (e.g., during and after the intervention), cancellation of appointments, or termination of services.
- Provides a copy of the signed and dated consent form at the conclusion of the meeting.

Adapted from *Ethical and Risk Management Issues in the Practice of ABA*, by E. Cipani, S. Robinson, and H. Toro, 2003. Paper presented at the annual conference of the Florida Association for Behavior Analysis, St. Petersburg, FL. Adapted with permission of the authors.

of a life-threatening emergency, or when a risk of serious harm is imminent. In selected jurisdictions, if a school district determines that certain services are necessary (e.g., special education programming), but the parents refuse, the district may have recourse through a process of progressive administrative review, mediation, and ultimately court action (Turnbull & Turnbull, 1998). As in all aspects of our work, providing treatment without consent requires the practitioner to consult and follow all applicable and current local, state, and national laws and regulations. These may differ across various jurisdictions and are subject to change over time.

Confidentiality

Professional relationships require **confidentiality**, meaning that any information regarding an individual receiving or having received services may not be discussed with, or otherwise made available to, any third party, unless that individual has provided explicit authorization for release of that information. Although confidentiality is a professional ethical standard for behavior analysts, it is also a legal requirement in many jurisdictions (Koocher & Keith-Spiegel, 1998). Figure 31.9 shows a standard release of information (ROI) form. Note that the ROI form clearly specifies what can be shared and released, and the expiration date for such releases.

Limits to Confidentiality and Legal Disclosures

Prior to the initiation of services, the behavior analyst must explain the limits of confidentiality to the client, document the meeting, and ensure that the client signs and dates the document. Disclosure of confidential information may occur without consent for a valid purpose, defined by the following conditions: (1) as mandated by law, (2) to comply with a court order, (3) when credible information suggests imminent harm or danger to the client or another person, (4) when an immediate crisis develops, or (5) when third-party payment for services is involved. For instance, if a student informs a school-based behavior analyst in confidence that another student has a loaded gun in the building, then confidentiality must be breached to protect the safety of students, staff, and administrators (cf. *Tarasoff v. Regents of the University of California, 1976*). Whenever such a disclosure is necessary, one should limit the disclosure to the absolute minimum information necessary to meet the requirement of the valid purpose.

Limits to disclosure also relate to third-party funding sources, which are entities other than the client or client guardian that pay for services. Typically, third-party entities have only limited access to confidential material pertaining to basic client information, the validation of the delivery of services, and payment for those services. Often, other treatment-specific information may only be provided to a third party if a formal release is signed by the "first-party" (i.e., the recipient of services or his or her guardian).

Maintaining Confidentiality

The practitioner must ensure that proper safeguards are in place to protect client confidentiality, including, but not limited to, storing hard-copy records in locked file cabinets, using only password-protected computer files, using encrypted files whenever possible, avoiding transmission of identifiable client information across any wireless systems, and confirming the identity of, and authorization for, any potential recipient of client-specific information.

Figure 31.9 Release of information form.

(Name of agency/provider)

RELEASE OF INFORMATION AND ASSESSMENT CONSENT FORM

CLIENT: ______________________________ DOB: __________

PARENT/GUARDIAN NAME: ______________________________

I consent for the above-named client to participate in an assessment through (provider/agency). I consent to have the assessment with the above-named individual conducted at the following locations (Circle relevant ones)

Home School Other: __________

I understand and consent to have the individuals responsible for care in the above-named locations involved in the assessment of the above-named client. In order to coordinate the assessment with these individuals, I authorize the release of the following confidential records to the individuals responsible for care in the above-named locations:

Evaluations/Assessments: ______________________________

IEP or Other Records: ______________________________

Other: ______________________________

I understand that these records may contain psychiatric and/or drug and alcohol information. I understand that these records may also contain references to blood-borne pathogens (e.g., HIV, AIDS). I understand that I may revoke this consent at any time; however, I cannot revoke consent for action that has already been taken. A copy of this release shall be as valid as the original. THIS CONSENT AUTOMATICALLY EXPIRES 30 DAYS AFTER TERMINATION OF SERVICES or after ONE CALENDAR YEAR.

PARENT/GUARDIAN: ______________________________ DATE: __________

BOX 31.1

When Use of Social Media Results in Breach of Confidentiality

Rebecca, a young BCBA, is excited that her 7-year-old client Jarod is making so much progress in acquiring language. She takes a selfie with Jarod and his family at a school picnic, and posts it on Instagram, with the family's verbal consent. Does this raise an ethical issue? Verbal consent is clearly not sufficient. Others may appear in the background of the photo, and have not given consent. Moreover, the family members might decide that they do not want to advertise to both friends and strangers that their child is the recipient of behavioral services (and thus a potential target for exploitation or abuse).

Louis is an aging BCBA from a state residential facility that closed 30 years ago, who now posts old pictures of staff and clients on a Twitter feed titled "Staff Reunion from State Institution X." Great pictures! Everyone loves to see them and share memories. Many of the individuals in these pictures have already passed away. Has Louis breached confidentiality here? The answer is almost certainly "yes." But is there real harm? What if Louis did a Google search on "images from State Institution X" and found pictures of his old clients already available on the Internet? Would posting those pictures on Twitter be an ethical breach? There is no way to evaluate the potential harm; but without a release from the client, this is an unauthorized breach of confidentiality.

The Health Insurance Portability and Accountability Act of 1996 (HIPAA), Public Law 104-191, specifically pertains to confidentiality of medical information and records. Many ABA-based services are delivered under the umbrella of medically authorized services (e.g., commercial insurance, Medicaid, other state funding). In these cases, practitioners must be careful to follow all HIPAA requirements, lest they sustain hefty fines and legal sanctions.

New and heretofore unforeseen challenges to confidentiality have arisen in recent years with the inexorable growth in the use of social media (see Box 31.1). Later in this chapter we will discuss issues relating to maintaining confidentiality in this expanding global landscape of public communication networks.

Breaches of Confidentiality

Breaches of confidentiality apply to two main response classes: (a) valid disclosures without consent and (b) unintentional breaches resulting from carelessness, misunderstanding of confidentiality requirements, or negligent disregard. For example, an unintentional and careless breach might occur if a behavior analyst provides confidential information to a parent about a child's status, without confirming that the parent requesting the information is the de facto legal guardian for the child. The behavior analyst should have been aware of who has authorized access, and release protected information only after confirming proper authorization.

Legally, all professionals must report suspected child abuse. Even in cases where the child fears repercussions if the abuse is reported, professionals at any level must step forward and report such cases to the authorities. Maintaining confidentiality when a child's health, safety, or welfare is involved does not apply. Figure 31.10 shows a sample form documenting the information provided to the client relating to abuse-reporting requirements.

Figure 31.10 Example of abuse-reporting protocol form.

(Name of agency/provider)

Confidentiality Act/Abuse-Reporting Protocol

Client: ____________________

I understand that all information related to the above-named client's assessment and treatment must be handled with strict confidentiality. No information related to the client, either verbal or written, will be released to other agencies or individuals without the express written consent of the client's legal guardian. By law, the rules of confidentiality do not hold under the following conditions:

1. If abuse or neglect of a minor, disabled, or elderly person is reported or suspected, the professional involved is required to report it to the Department of Children and Families for investigation.
2. If, during the course of services, the professional involved receives information that someone's life is in danger, that professional has a duty to warn the potential victim.
3. If our records, our subcontractor records or staff testimony are subpoenaed by court order, we are required to produce requested information or appear in court to answer questions regarding the client.

This consent expires 1 year after signature date below.

____________________ ____________________

Parent/Legal Guardian Date

Figure 31.11 Defining an acceptable treatment environment.

1. The environment is engaging: Reinforcement is readily available; problem behaviors are reduced; reinforcement contingencies are effective; exploratory play and practice flourish; the environment is by definition humane.
2. Functional skills are taught and maintained: Environments are judged not by paperwork or records, but by observed evidence of training and progress; incidental teaching allows the natural environment to support skill acquisition and maintenance.
3. Behavior problems are ameliorated: A purely functional approach regarding both behaviors and procedures leads to effective interventions; arbitrary labels based on topographies are replaced by individualized definitions based on function.
4. The environment is the least restrictive alternative. Again, this is functionally defined, based on freedom of movement and activity engagement parameters; community settings may actually be more restrictive than institutional ones, depending on their behavioral effects.
5. The environment is stable. Changes in schedules, programs, peers, caregivers, etc., are minimized; skills are maintained; consistency and predictability valued.
6. The environment is safe: Physical safety is paramount; supervision and monitoring are adequate; peer review ensures that proper program procedures are based on function.
7. The client chooses to live there: Efforts are made to determine client choice; alternative environments are sampled.

Protecting the Client's Dignity, Health, and Safety

Dignity, health, and safety issues often center on the contingencies and physical structures present in the environments in which people live and work. The behavior analyst should be acutely aware of these issues. Favell and McGimsey (1993) provided a list of acceptable characteristics of treatment environments to ensure dignity, health, and safety (see Figure 31.11).

Dignity can be examined by addressing the following questions: Do I honor the person's choices? Do I provide adequate space for privacy? Do I look beyond the person's disability and treat the person with respect? Behavior analysts can help to ensure dignity for their clients by defining their own roles. By using operant principles of behavior, they teach skills that will enable learners to establish increasingly effective control over the contingencies in their own natural environments. Everyone has the right to say yes, no, or sometimes nothing at all (cf. Bannerman, Sheldon, Sherman, & Harchik, 1990).

Choice is a central principle in the delivery of ethical behavioral services. In behavioral terms, the act of making a choice requires that both behavioral and stimulus alternatives be possible and available (Hayes et al., 1994). Practitioners must provide the client with behavioral alternatives, and the person must be capable of performing the actions required by the alternatives. To leave a room, a person must be physically capable of opening the door, and the door must not be blocked or locked. The term *stimulus alternatives* refers to the simultaneous presence of more than one stimulus item from which to choose (e.g., choosing to eat an apple instead of an orange or pear). To have a fair choice, a client must have alternatives, must be able to perform each alternative, and must be able to experience the natural consequences of the chosen alternative.

Negligence represents a failure to demonstrate professional integrity on the part of the provider, and threatens the client's health and safety [see *The Compliance Code* 1.04(c)]. Negligence is typically exhibited as nonfeasance—not doing what ought to be done. For instance, a behavior analyst may take on too many cases, and not adequately monitor progress for those cases; or he may not have planned for adequate clinical coverage during his vacation, leaving clients without access to services during those weeks; or perhaps she abruptly terminates a case without consideration for the client's need for a slower fade-out of services. Additionally, negligence can sometimes be demonstrated when acting in an imprudent fashion. For example, an inexperienced analyst might escort too many clients on a community outing, and lacking appropriate supervision, one of the clients wanders away. This was not the intention of the escort, but by not providing adequate supervision, a negligent act occurred.

While negligence is typically unintentional, **fraudulent conduct** is characterized as intentional, willful, and deceitful behavior, and can cause harm to another individual. Misrepresenting one's credentials, training, or background to a recipient of services is a form of fraud, as is overbilling or falsifying client records. Less obvious forms of fraudulent conduct, such as minor falsifications of raw data (e.g., a data collector fills in a data sheet well after the observational period has ended) may or may not result in formal sanctions, but still constitute a clear breach of professional ethics. When an investigation leads to a finding of negligence or fraudulent conduct on the part of a practitioner, professional sanctions are nearly certain, and legal action resulting in fines or even prison is possible, depending on the circumstances and the degree of demonstrated harm.

COORDINATING WITH OTHER PROFESSIONALS

Behavior analysts are often required to work with a myriad of other professionals, many of whom subscribe to scientifically supported treatments. Such collaborations among interdisciplinary team members can result in a variety of positive educational and behavioral outcomes for individuals with disabilities (Brodhead, 2015; Hunt, Soto, Maier, & Doering, 2003) and promote treatment fidelity as well (Kelly & Tincani, 2013).

For example, behavior analysts often coordinate with medical professionals, particularly when the client is receiving psychotropic medication to address problem behavior. Behavior analysts must be sensitive to the fact that the medical professionals follow their own professional ethical code, as does each professional member of any treatment team. Effective coordination between disciplines can only occur if communication among all treatment team members remains clear and frequent, particularly in relation to any proposed "changes in medication regimens, therapies, or programming." (Newhouse-Oisten, Peck, Conway, & Frieder, 2017, p. 148).

Behavior analysts may also interact with staunch proponents of nonbehavioral treatments for which solid evidence of effectiveness is unknown, ignored, or unreliable and therefore invalid. In these situations, behavior analysts must call upon their technical skills, social interaction skills, and *The Compliance Code* when formulating a comprehensive plan for the client. *The Compliance Code* [2.09(d)] states, in part, that behavior analysts have a responsibility to "review and appraise the effects of any treatments about which they are aware that might impact the goals of the behavior-change program" (BACB, 2014, p. 9). Newhouse-Oisten et al. (2017) suggest that behavior analysts must focus on two aspects of each suggested intervention: "whether or not the intervention is supported by research evidence, and whether or not the intervention change is compatible with other interventions" (p. 150). This type of analysis is a clear ethical imperative when determining best practice, but can sometimes cause conflict within a treatment team. Brodhead (2015) provides a series of yes-no questions in a flow chart for evaluating such unknown treatments in an unbiased and objective manner: (1) What is the nonbehavioral treatment? (2) Is the client's safety at risk? (3) Are you familiar with the treatment? (4) Being now familiar with the treatment, have new safety concerns been revealed? (5) Is treatment success possible when nonbehavioral treatment is translated into behavioral processes? (6) Are the impacts to the client sufficient to justify the possibility of compromising the professional relationship? Brodhead also provides a checklist for analyzing proposed treatments and suggests options for ethical action by behavior analysts.

An alternative, but similar, approach is provided by Schreck and Miller (2010), who provide a flow chart to help the practitioner evaluate the validity of three aspects of the proposed but unknown treatment: (1) the underlying theory, (2) the treatment techniques, and (3) the claims of effectiveness.

Box 31.2 offers additional guidelines for professionals to help parents, teachers, and others to evaluate alternative treatment recommendations that may arise.

The best interests of the client will be served when the behavior analyst uses one of these models, or an alternative effective and structured approach, to help guide collaboration with other professionals in a way that enhances client services and reflects well on the field. However, in rare cases, the alternative approach is so disruptive to the effective delivery of behavioral services that the practitioner must withdraw from the case, providing the client with clear and valid reasons for doing so.

This approach to coordinating with other professionals can be applied to the third scenario described in the opening of the chapter. Ms. Dougherty is in a bind, but she only "perceives" the intent of the school administrator with whom she has to coordinate services. She should continue to advocate for the best interests of

BOX 31.2

Guidelines for Professionals to Help Parents, Teachers, and Other Caretakers

Parents, teachers, and other caretakers are not without resources when trying to ascertain whether a particular treatment meets the standard for best practice, or falls within the scope of a fad or untested intervention. Three resources follow.

First, change agents can consult a trusted behavior analyst and the corresponding Code of Conduct that he or she follows. *The Compliance Code* states that, "Behavior analysts always have the obligation to advocate for and educate the client about scientifically supported, most effective treatment procedures" [*The Compliance Code* 2.09(a)]. Moreover, *The Compliance Code* 6.02 states: "Behavior analysts promote behavior analysis by making information about it available to the public through presentations, discussions, and other media." So, when change agents follow the advice of a behavior analyst, they increase the chances that the recommendations they receive meet quality care criteria.

Second, Schreck and Mazur (2008) suggest a series of six strategic techniques to counter the proliferation of fad/alternative treatments: (1) compare the effectiveness of the treatment in question with ABA; (2) emphasize the long-term cost effectiveness of ABA; (3) publicize results of these comparisons via mass media; (4) market ABA to parents and employers; (5) find a way to short-circuit the search for an "easy way out" solution and the influence on this by mass media; and (6) educate BCBAs on the ethical dilemmas and consequences of using nonsupported treatments.

Finally, Heron, Tincani, Peterson, and Miller (2005) suggest that all interventionists develop what noted astronomer, Carl Sagan, termed a "Baloney Detection Kit"—a kind of toolbox of conceptual instruments to promote skeptical thinking and logical analysis. Peters and Heron (1993) add that practitioners, when trying to judge the efficacy of any educational or therapeutic approach, should consider answering these five questions:

1. Does the model or program derive from a sound theoretical base?
2. Is the research methodology convincing and compelling?
3. Is there consensus with existing literature?
4. Is there evidence that outcome data are consistently produced?
5. Is there evidence of social validity?

the student, but also should consult with her supervisor to assist with her advocacy, and suggest a systematic assessment of the suggested treatments to identify the most effective approach.

When a behavior analyst makes a referral to another professional, there must never be referral fees or kick-backs. The practitioner should provide the client with a list of options, a choice of possible providers, and allow the client to make the final choice. No money, gifts, or favors should ever change hands when a referral from one professional results in a client contacting another professional.

SOCIAL MEDIA AND NEW TECHNOLOGIES

The use of social media is pervasive, with compelling evidence to show that a great majority of the world's population has access to, and frequently uses, smartphones, tablets, pads, and other electronic devices to share information and interact with others (Greenwood, Perrin, & Duggan, 2016; Pagoto et al., 2016). In effect, social media and the Internet have become the marketplace and village square of the 21st century, where social interactions flourish. Admittedly, social media present positive opportunities for sharing information and discussing issues pertaining to behavioral science (e.g., *Students of Behavior Analysis* on Facebook). Several media platforms (e.g., Facebook, Twitter) offer private areas where postings may be viewed by an individual, or a select few, who are provided specific and limited access. Thus, a practitioner might send a private message to an individual or limited group, to assist in supporting treatment goals or to help maintain treatment fidelity (e.g., sending reminders to collect data).

Moreover, new technologies provide individuals with the opportunity to view treatment sessions from remote locations, in real time, which can be useful for supervisors, parent/guardians, and behavior analysts in training.

Still, all therapeutic disciplines must come to terms with the proper way to use social media, without breaching confidentiality or falling into ethical quicksand. Unsecured transmission networks are of significant concern, and behavior analysts must be well versed in the nature and scope of the security provided to protect their communications. Even in supposedly secure social media sites, users must be cautious, since data breaches are always possible and security settings are subject to change by the service provider. Thus, client-specific information must always be masked, such that any individual discussed can never be identified by name or location (Pagoto et al., 2016). While codes of professional ethics are only recently beginning to catch up with the challenges posed by these new media, O'Leary, Miller, Olive, and Kelly (2017) have provided useful recommendations for maintaining professional ethics in an emerging social media environment (see Box 31.3). Further, Box 31.4 presents ways to avoid the pitfalls of using social media.

Helping the Client Select Outcomes and Behavior Change Targets

The recipient of behavior analysis services (or the recipient's guardian) should identify the ultimate goal(s) of the service, termed "outcomes." Outcomes must be meaningful to the individual, contextually appropriate, and achievable. The goals should focus on a quality of life issue for the individual (e.g., obtaining employment, establishing long-term relationships, accessing community and educational settings) (Felce & Emerson, 2000; Risley, 1996). The behavior analyst should help the individual express her preferences and life goals, and identify what behaviors she would need to do (and not do) to help accomplish those goals. Behaviors selected for change must benefit the individual recipient of services, not the practitioner or caregiver.

Early on, behavior analysts were rightly criticized for focusing on behaviors that mainly benefited staff (e.g., Winett &

BOX 31.3

Recommendations for Behavior Analysts' Use of Social Media

1. *Real clients should be heavily disguised.* Whenever talking about anything client specific, many details can be changed and the point of the discussion remains intact.
2. *Avoid making treatment recommendations.* There is simply not enough information on social media outlets to give specific advice.
3. *Refer readers back to the literature.* Rather than making a treatment recommendation, recommend a specific paper, or work by particular authors.
4. *Write a disclaimer.* Be sure that anyone reading anything you write, especially an answer to a question, cannot get the idea that there is a professional relationship. Be explicit in stating this.
5. *Provide resources.* Beyond literature references, provide interested parties with ideas for search topics, websites, and the opportunity to discuss a post offline, to prevent misinterpretation by non–behavior analysts.
6. *Provide organizational training.* This fits with *The Compliance Code's* requirement that we create a culture of ethics. Training in a work setting on how to navigate the dangerous currents of social media can be vital in preventing potential problems.

From "Blurred Lines: Ethical Implications of Social Media for Behavior Analysts," by P. N. O'Leary, M. M. Miller, M. L. Olive, and A. N. Kelly, 2017, *Behavior Analysis in Practice, 10*, pp. 45–51. Used with permission.

BOX 31.4

The Pitfalls of Social Media

Guy is a veteran behavior analyst who reads young BCBAs' case-related questions on a Facebook page titled "Behavior Analysis Q&A!" and provides answers and treatment advice to help guide them along. Everyone on the Facebook page tells Guy how much his help is appreciated. Even if the posted questions are carefully written so they contain no hints of client identity or specific setting, are there ethical issues involved here? Yes, many. The young practitioners might begin to rely on the sage advice provided by this "Facebook expert" and possibly others who provide quick answers. Those posing questions may, over time, become less likely to search the literature, seek necessary supervision or mentoring, or obtain consultations. They may eventually engage in the same form of "social media mentoring," since Guy has modeled this as a form of appropriate professional behavior. Any advice given on a case-related question, no matter how wise it may be, is based on incomplete information. Others reading these posts, including non–behavior analysts, might misuse these treatment suggestions by inappropriately applying them to other individuals with other behavior problems in other settings. Everything about Guy's apparently supportive action violates the core focus of our clinical approach: to directly evaluate the behavior of the individual in context, and base treatment on identified functional relations. The behavior of this seasoned adviser may help a few people, but he will almost certainly never know the extent of the damage he may well have caused.

Winkler, 1972). For example, an adult in a work setting might easily have had docile "compliance" identified as a target behavior, even though strengthening work skills and functional social interactions would have been of far greater benefit to that individual. Compliance, once commonly identified in behavior programs, is not an appropriate goal by itself. The question must be posed: compliant to what? Teaching compliance is ethically justifiable only when it facilitates the development of other functional or social skills that will, in turn, contribute to that individual's obtaining long-term self-identified outcomes.

This discussion on appropriate selection of goals, along with prior information on obtaining input from a guardian, surrogate, or advocate may be applied to the first scenario described in the opening of the chapter. Dr. Han could suggest that Lee meet with a trusted adviser or surrogate to review the pros and cons of the move from a residential program to an apartment. In addition, they could discuss the reasons Lee wants to move, to see if the current setting could be adjusted in ways that would meet his needs more effectively, without having to go through a move in residence.

Maintaining Records

Well-maintained records (e.g., assessments, treatment plans, progress notes, data) facilitate the provision of future services, meet agency and/or institutional requirements, ensure accurate billing, allow for future research, and help meet legal requirements. Behavior analysts must typically maintain all case file records for 7 years, subject to individual, state, or jurisdictional requirements. *The Compliance Code* states that the behavior analyst is responsible for maintaining individual client confidentiality in terms of "creating, storing, accessing, transferring, or disposing of records under their control, whether these are written, automated, electronic, or in any other medium" [*The Compliance Code* 2.07(a)]. When destroying paper records, use crosscut shredding. Destruction of electronic records requires that all stored files in all possible storage locations be fully expunged. Be aware that electronic transmission of confidential records across any unsecured medium (e.g., fax lines in public areas, e-mail) is prohibited under Public Law 104-191 (HIPAA).

ADVOCATING FOR THE CLIENT

Providing Necessary and Needed Services

Prior to initiating services, a behavior analyst has the responsibility to validate that a referral warrants further action. This poses the first ethical challenge to the practitioner: deciding whether to accept or reject the case. The decision to provide treatment may be divided into two sequential decisions: (1) Is the presenting problem amenable to behavioral intervention? and (2) Is the proposed intervention likely to be successful?

Is the Problem Amenable to Behavior Treatment?

To determine whether behavioral intervention is necessary and appropriate, the behavior analyst should seek answers to the following questions:

1. Has the problem emerged suddenly?
   a. Might the problem have a medical cause?
   b. Has a medical evaluation been done?
2. Is the problem with the client or with someone else? (For example, a student who has done well through fourth grade suddenly exhibits "behavior problems" in fifth grade, although she remains well behaved at home. Perhaps a simple change of teachers will solve the problem.)
3. Have other interventions been tried?

4. Does the problem actually exist? (For example, a parent has serious concerns about a 3-year-old child who "refuses to eat" everything provided at every meal.)
5. Can the problem be solved simply or informally? (For example, the same 3-year-old "elopes" through a back door that is left wide open all day.)
6. Might the problem be better addressed by another discipline? (For example, a child with cerebral palsy may need adaptive equipment instead of behavioral treatment.)
7. Is the behavioral problem considered an emergency?

Is the Proposed Intervention Likely to Be Successful?

Questions to ask when considering whether the intervention is likely to be successful include the following:

1. Is the client willing to participate?
2. Are the caregivers surrounding the client willing or able to participate?
3. Has the behavior been successfully treated in the research literature?
4. Is public support likely?
5. Does the behavior analyst have the appropriate experience to deal with the problem?
6. Will those most likely to be involved in implementing the program have adequate control of the critical environmental contingencies?

If the answer to all six of the questions posed in this second list is "yes," then the behavior analyst can act. If the answer to any of these six questions is no, the behavior analyst should seriously consider declining to initiate treatment.

Embracing the Scientific Method

General questions related to what constitutes science and how it applies to the study and improvement of behavior were addressed in Chapter 1. The scientific method relates to ethics, at least in part, in its insistence that evidence-based practices utilize direct measurement to establish the effectiveness of methods that are grounded in peer-reviewed literature, and tests any emerging practice to assess its effectiveness before implementation. The importance of embracing evidence-based practices increases as questionable claims regarding effective "alternative" treatments become more widespread, and are seemingly accepted by the general public without critical examination (Shermer, 1997). In Sagan's (1996) view, extraordinary claims require extraordinary evidence, and as space engineer James Oberg wryly quipped, "In science keeping an open mind is a virtue, but not so open that your brains fall out" (cited in Sagan, 1996, p. 187).

Applied behavior analysts should base their practice on two primary sources: scientific literature and frequent direct measurement of behavior in context. Peer-reviewed scientific reports published in reputable journals and well-referenced texts provide objective information on effective intervention strategies. Direct repeated measurement of behavior provides empirical data to inform our assessments and evaluations of the effectiveness of our work. The cornerstone of ethical practice is providing effective services based on solid, replicated research evidence.

Alternative/Fad Treatments

The educational and therapeutic world abounds in a wide variety of teaching techniques and approaches to address problem behavior. Most of these so-called remedies have never been scientifically or empirically validated. As new, untested treatments arise, parents, teachers, and other caregivers who legitimately seek educational, social, or behavioral assistance can easily fall prey to ill-advised and uninformed pied pipers (Heron et al., 2005). For example, Kohn (1997) promotes the notion that repeated practice of a skill is harmful because it destroys a student's motivation to learn, and ultimately results in the student being unhappy with school in general. This assertion has gained traction with the general public despite a complete lack of peer-reviewed evidence to validate the claim. On the contrary, decades of carefully controlled and replicated research have demonstrated that repeated practice is a key determinant in the mastery of a skill (Binder, 1996; Ericsson & Charness, 1994).

Box 31.5 explores several issues and problems related to the continuing emergence of alternative and fad treatments.

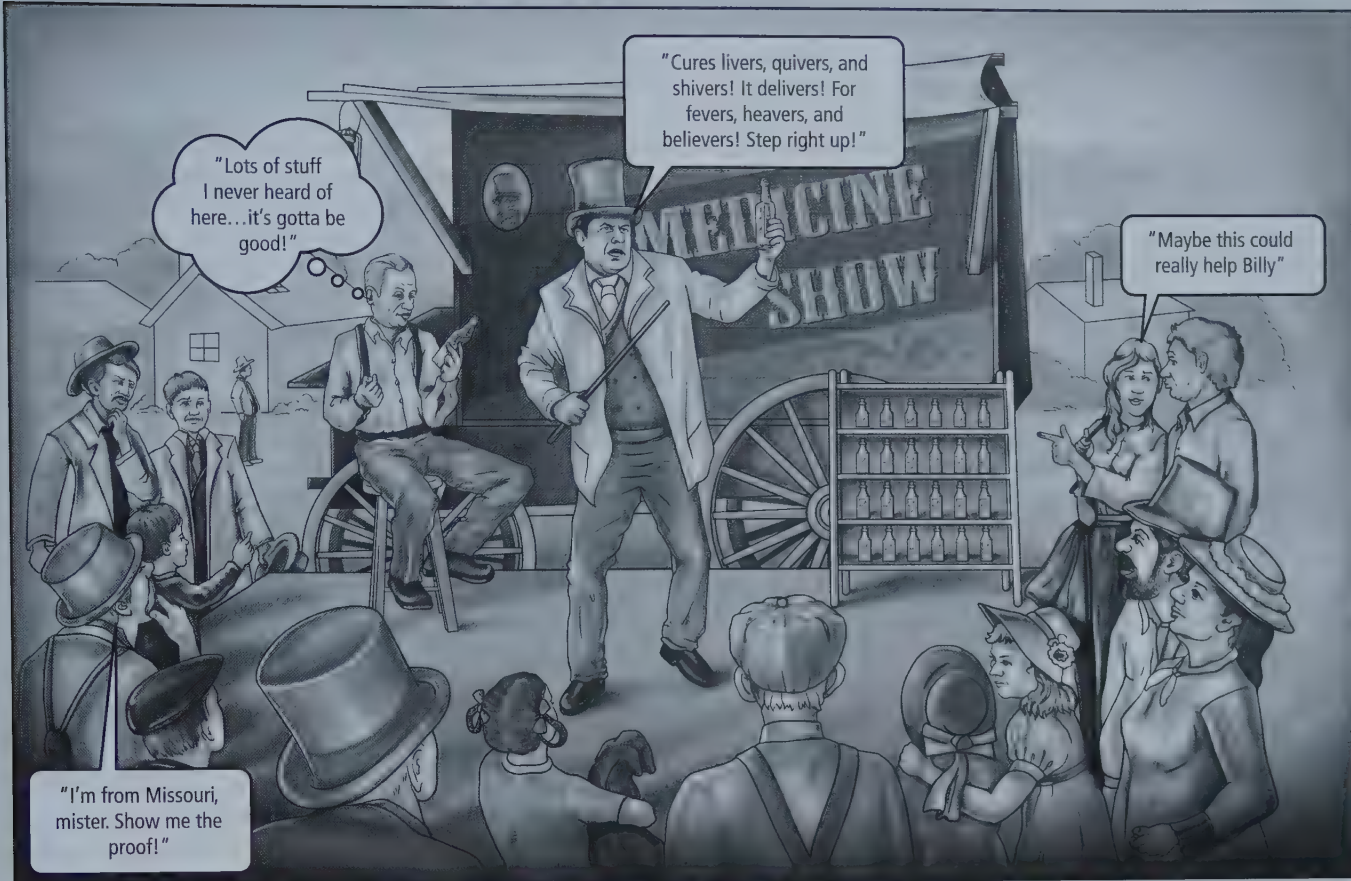
Evidenced-based Best Practice and Least Restrictive Alternatives

Silvestri and Heward (2016) define evidence as "the outcome of the application of the scientific method to test the effectiveness of a claim, a theory, or a practice." Moreover, they state that, "extraordinary evidence requires replication. A single study, anecdote, or theoretical article, no matter how impressive the findings or how complicated the writing, is not a basis for practice" and that "an extraordinary claim based on anything but proportionally strong evidence should bring skepticism" (pp. 149–150).

An essential component of ethically driven behavior analysis is not only that interventions and related practices be evidence-based, but also that the most powerful, yet least intrusive methods be used first. Moreover, intervention plans must be designed, implemented, and evaluated systematically. If the individual fails to show progress, data systems should be reviewed and, if necessary, interventions modified. If progress is made, the intervention should be phased out and assessed for generalization and maintenance. In all phases of an intervention, data and direct observations should drive treatment decisions. Models and programs that meet the above standard are deemed "best practice" and can be judged by their meeting at least five standards: Does the model or program derive from a sound theoretical base? Is the research methodology convincing and compelling? Is there consensus with existing literature? Is there evidence that outcome data are consistently produced? Is there evidence of social validity? (Peters and Heron, 1993). *The Compliance Code* [2.09(c)] also indicates that when "more than one scientifically supported treatment

BOX 31.5

Lotions, Potions, and False Notions



The promotion and sale of ineffective treatments for a variety of ailments has a long history. "Snake-oil salesmen" were so common in the Wild West era of the United States that the term is now used to label any purveyor of fake cures or, more broadly, false solutions to problems. Individuals who are likely to seek behavior analytic services have often been confronted with a dizzying array of claims and counter-claims regarding a wide variety of alternative treatment options. Many promise spectacularly positive results and even total "cures" for everything from confrontational teenage behavior to the full spectrum of autism-related disorders. Most of these claims have no solid foundation in research or peer-reviewed literature. Treatments such as *DIRFloortime*, *Doman-Delacato patterning treatment*, *aromatherapy*, *animal-assisted therapies*, and *auditory integration training* (to name but a few) lack reputable scientific evidence base (Hines, 2001; Koocher & Gill, 2016; Mudford & Cullen, 2016). Other alt-treatments have been extensively studied and soundly debunked, as in the case of *facilitated communication* (Jacobsen, Foxx, & Mulick, 2016; Lilienfeld, Marshall, Todd, & Shane, 2015). We expect other untested treatments to enter the marketplace, many likely to use new technology-based interventions that sound highly scientific but lack a firm basis in research and replication. When a behavior analyst is providing services to an individual who may be tempted to try one or more of these alt-treatments, he or she has a responsibility to assist that individual in evaluating the likely effectiveness and potential costs related to the proposed intervention.

has been established, additional factors may be considered in selecting interventions, including, but not limited to, efficiency and cost-effectiveness, risks and side-effects of the interventions, client preference, and practitioner experience and training" (2014, p. 9). Geiger, Carr, and LeBlanc (2010) describe a similar approach to selecting among function-based treatments for problem behavior, taking into account consideration of implications for safety, the resources of the environment, and prioritization of improved quality of life and skills for the client (see Figure 31.12).

Figure 31.12 Selecting function-based interventions.

Presenting Concern	Ethical Problem-solving Strategies
Background: Marco is a 3-year-old child diagnosed with autism spectrum disorder. He has very minimal verbal skills but is fortunate enough to be enrolled in an early intensive behavioral intervention program. He is making progress in developing new skills more rapidly than others. He often has tantrums with his parents, particularly when they are performing activities of daily living (e.g., dressing him) and hygiene maintenance (e.g., brushing his teeth). During these tantrums, he yells, cries, and engages in self-injurious behavior (e.g., scratching at his own arms) as well as assaultive behavior (e.g., scratching and pinching others). Marco has begun to exhibit these same behaviors with the staff who work with him in his behavioral program. His BCBA has conducted a comprehensive functional assessment, including a functional analysis. The identified function of these behaviors is escape (i.e., negative reinforcement function). The BCBA must now develop a behavior plan to reduce his problem behavior and increase his appropriate completion of tasks. This BCBA has read a number of articles from the literature and discovered that many evidence-based interventions are function-based treatments for escape-maintained problem behavior, and he is unsure which of them would be best suited to this situation.	*Analysis and Action*: Marco's BCBA identified a published clinical decision-making model for selecting function-based treatments for escape-maintained problem behavior (Geiger, Carr & LeBlanc, 2010). This model suggests that five core considerations should guide the determination of which intervention is best suited for a specific situation: Quality of life of the client, safety of everyone involved, resource constraints of the environment, existing skill sets of the client, and remaining skill needs of the client. The BCBA answered the progressive list of questions suggested in this model and determined that the optimal intervention would be functional communication training (FCT). The tasks that are occurring now are necessary and the instructional procedures are generally quite good. Marco's staff and parents felt as though they could remain safe and tolerate a small amount of problem behavior as long as it is decreasing over time. Marco will comply with several demands, and it is acceptable to allow periodic breaks from tasks as part of a treatment plan. This allowed Marco's BCBA to consider the most important skill to prioritize. Since Marco has very limited verbal skills, it is important to teach him a way to ask for a break or assistance when he finds himself in an aversive situation. So FCT is a great first intervention, and other interventions might be added if needed.

CONFLICT OF INTEREST

A **conflict of interest** occurs when a principal party, alone or in connection with family, friends, or associates, has a vested interest in the outcome of the interaction. The most common form of conflict arises in the form of dual- or multiple-role relationships. Conflicts arise when a person acting as a therapist enters into another type of relationship with the client, a family member, or a close associate of the client, or promises to enter into such a relationship in the future. These relationships may be financial, personal, professional (providing another service, for example), or otherwise beneficial to the therapist.

Direct and frequent observations during assessment and intervention phases bring the behavior analyst into close contact with the client (and often family members, other professionals, and caregivers) in various natural settings. Personal relationships can be formed that may imperceptibly cross professional boundaries. For example, family members may offer unsolicited gifts or make invitations to parties or other events. In every interaction related to their professional duties, behavior analysts must monitor their own behavior with vigilance, guarding against blurring the lines between personal and professional conduct. This is especially true when treatment is provided within a private home. In service contexts, personal relationships of any kind can quickly develop ethical complications, and are to be avoided.

Other professional conflicts of interest might arise as well. For example, a teacher should never hire a current student as an employee in a business outside of school. That student's performance in one area could potentially affect the teacher–student interaction in the other. A supervisee should be off limits as a potential romantic partner. A member of a peer review committee must not participate in a review of his own work, or the work of his supervisees. The general rule to follow is that the practitioner should strive to avoid all potential conflicts of interest. A practitioner in doubt should consult with a supervisor or trusted and experienced confidant.

CREATING A CULTURE OF ETHICAL PRACTICE

Ethical or unethical behavior occurs in an environmental context, and environments can be arranged to support behavior that is ethical. Moreover, "the ability to sort through the issues and make decisions that most people would see as ethical is a skill that emphatically can and should be taught" (Lattal & Clark, 2007, p.13). Even the most proficient professionals benefit from robust environmental supports for problem-solving repertoires that help maintain ongoing ethical practice. To foster a culture of ethical practice (*The Compliance Code* 7.01), organizations can commit to providing ethics-related trainings, support systems, and differential reinforcement of ethical behavior. Brodhead and Higbee (2012) outline various strategies for teaching and maintaining ethical behavior in a professional organization. They address the importance of initial training in ethics and ongoing training and supervision

BOX 31.6

Does Your ABA Organization Have an Ethics Director?

The Ethics Network at Linda LeBlanc's clinical service organization includes a position of Ethics Director. This doctoral-level behavior analyst has extensive experience as a supervisor and extensive training in ethics. The Ethics Director serves as a mentor to the ethics and assistant ethics chairpersons, each of whom fosters and supports ethics throughout the organization. The Ethics Network engages in four primary activities designed to sustain continuous discussion about ethics at all levels of the organization.

First, the network director designed the initial ethics training for all members of the organization with a focus on (1) core principles that underlie all aspects of our ethical code and (2) structured practice in a six-step problem-solving process that includes detection and (3) an analysis of situations as potential ethical dilemmas.

Second, the network director monitors an internal ethics hotline that can be accessed by any member of the organization who needs support with a difficult or ambiguous situation. Asking for assistance through this mechanism is differentially reinforced with swift assistance and praise for critical analysis of the situation, as prompted by the submission portal (i.e., the problem-solving steps). The ethics leadership team analyzes topics and situations submitted to the hotline to hone team members' skills in responding to ethical conundrums. Those situations are used monthly to create instructional materials called "discussion starters"; these include specific role plays that can be used in supervision activities throughout the organization.

Third, semi-annual trainings occur with the clinical team to review material that is pertinent to the most challenging situations submitted to the hotline (e.g., ethical issues associated with serving families who are experiencing divorce or separation).

Fourth, the ethics network director leads discussions with the administrative leaders of the organization at least annually to focus on specific parts of *The Compliance Code* that are directly pertinent to their ongoing activities (e.g., review of codes relevant to social media and nonsolicitation of testimonials with the marketing team). Since these individuals generally have very different training and backgrounds than behavior analysts, they may be likely to misrepresent important features of our code and our field unless the behavior analyst accepts the responsibility to educate them about the underlying principles that lead us to value certain actions (Code 10.07).

Establishing the Ethics Network has been personally and professionally rewarding, as it has led to improved communication, enhanced problem solving, team collaboration, and higher quality service to our clients.

that directly incorporates ethics as a professional development topic. In addition, they recommend using previous ethically complex situations as instructional opportunities to prevent recurrence of the same problem.

In sum, continuous education and discussion about ethics and the contingencies that support analysis and assistance seeking are core features of an environment designed to maintain ethical behavior in an organization (see Box 31.6). When individual behavior analysts do not operate within an environment with other behavior analysts, they should initiate connections with other behavior analysts through professional organizations, conference sessions, or seminars to access support and consultation as needed.

CONCLUSION

Practicing ethically poses daily challenges. It requires vigilance, self-monitoring, and the ongoing dynamic application of one's principles and professional code. Our profession has created a code of ethical conduct and set consequences for violating those rules. But we can also create cultures of ethics in a wide variety of settings, to help train practitioners in ethical decision making and to support ongoing discussion of ethical issues. In this way, we take an antecedent approach that promotes ethical action and effectively prepares the behavior analyst for acting ethically at the moment when an ethical dilemma presents itself, the moment when unethical behavior is most likely to occur.

Most ethical challenges, regardless of the details involved, may best be addressed by revisiting the three questions posed early in the chapter: What is the right thing to do? What is worth doing? What does it mean to be a good behavior analyst? These three questions can serve as focal points for effective and ethical decision making. Addressing these questions honestly and without bias will help practitioners avoid many potential ethical pitfalls and keep their efforts centered on the interests of the client, the student, or others receiving their services.

By practicing ethically, and by maintaining adherence to the scientific method and the principles, procedures, and dimensions of applied behavior analysis, practitioners have a ready source of valid, accurate, reliable, and believable data to inform decision making. As a result, they will be more likely to achieve the original and continuing promise of applied behavior analysis: to apply knowledge gained through research in both the experimental analysis of behavior and applied behavior analysis to help solve problems of social importance to each individual recipient of services.

SUMMARY

What Is Ethics and Why Is It Important?

1. For the applied behavior analyst, ethics defines specific rules of conduct.
2. Two approaches to ethics are termed the *deontological* or *"acts-oriented" approach*, which evaluates only the act itself, and the *utilitarian* or *"results-oriented" approach*, which looks at the results or consequences of the act. Both approaches contribute to the ethical conduct of behavior analysts.
3. Ethics describes behaviors, practices, and decisions that address three basic and fundamental questions: What is the right thing to do? What is worth doing? What does it mean to be a good behavior analyst?
4. Ethics is important because it helps practitioners decide whether a course of action related to the professional practice of behavior analysis is appropriate or inappropriate. Codes of ethics guide decisions on how to act even in the presence of outside pressures, priorities, or expediencies.
5. Ethical practices are derived from other behavior analysts and professionals to ensure the well-being of (a) clients, (b) the profession itself, and (c) the culture as a whole. Over time, practices become codified into ethical rules of conduct that may change in response to a changing world.
6. Personal histories, including cultural and religious experiences, influence how practitioners decide on a course of action in any given situation.
7. Practicing behavior analysts must be aware of the settings in which they work and the specific rules and ethical standards applicable to those settings. Behavior analysts are required to be fully familiar with all ethical and legal issues pertaining to their practice.

Standards of Professional Practice for Applied Behavior Analysts

8. Professional organizations adopt formalized statements of conduct, codes of professional behavior, or ethical standards of behavior to guide their members in making decisions. Also, standards exact sanctions when deviations from *The Compliance Code* occur.
9. The client's welfare is at the forefront of ethical decision making.
10. Behavior analysts must take steps to ensure that their professional conduct is self-regulating.
11. Professional standards, guidelines, or rules of practice are written statements that provide direction for conducting the practices associated with an organization.
12. Behavior analysts are guided by position statements and ethical codes such as the *Professional and Ethical Compliance Codes for Behavior Analysts, Ethical Principles of Psychologists and Code of Conduct, The Right to Effective Behavioral Treatment, Statement on Restraint and Seclusion*, and *The Right to Effective Education.*
13. Behavior analysts are now held to *The Compliance Code*, and perceived violations of that code can be reported using a notice of alleged violation.

Ensuring Professional Competence

14. Professional competence in applied behavior analysis is achieved through formal academic training that involves coursework, supervised practica, and mentored professional experience.
15. Behavior analysts must be truthful and accurate in all professional and personal interactions.
16. Competence in supervision is critical, as most behavior analysts supervise the work of others in the context of their practice.
17. Formal certification and/or licensure credentials in behavior analysis allow consumers to identify practitioners who have demonstrated at least a minimum standard of training and competence.
18. Behavior analysts must practice within their areas of competence, training, and experience.
19. Behavior analysts can develop additional competencies through supervised or mentored experience.
20. Behavior analysts take responsibility for all facets of the undertaking when providing supervision and/or mentoring of other behavior analysts, individuals seeking certification in behavior analysis, paraprofessionals, or non–behavior analysts.
21. Behavior analysts have a responsibility to maintain and expand professional competencies, and may do so using continuing education, participation in professional conferences, maintaining contact with current professional literature, and/or participating in peer review committees.

Ethical Issues in Client Services

22. Three tests must be met before informed consent can be considered valid: capacity to decide, voluntary decision, and knowledge of the treatment.
23. Confidentiality refers to the professional standard requiring that the behavior analyst not discuss or otherwise release information regarding someone in his or her care. Information may be released only with formal permission of the individual or the individual's guardian.

24. Practicing behavior analysts have a responsibility to protect a client's dignity, health, and safety. Various rights must be observed and protected, including the right to make choices, the right to privacy, the right to a therapeutic treatment environment, and the right to refuse treatment.

25. Negligence (not doing what one should) and fraudulent conduct (willful deceit likely to cause harm) on the part of the behavior analyst pose a direct threat to the client's health and safety, and can result in professional sanctions and possibly legal penalties.

Coordinating with Other Professionals

26. Behavior analysts must coordinate services with other professionals in a way that best meets the interests of the individuals receiving services, and that represents the field of behavior analysis in a positive and professional manner.

27. Effective coordination between disciplines can only occur if communication among all treatment team members remains clear and frequent.

28. The best interests of the client will be served when the behavior analyst guides collaboration with other professionals in a way that enhances client services, and reflects well on the field.

Social Media and New Technologies

29. The use of social media and the emergence of other new technologies present an increasing array of opportunities, challenges, and potential ethical pitfalls for the behavior analyst. Special care must be taken to protect client confidentiality, and the behavior analyst should avoid making case-specific treatment recommendations in a public forum.

30. New technologies allow practitioners to view treatment sessions from remote locations, in real time, which can be useful for supervisors, parents/guardians, and behavior analysts in training.

31. All therapeutic disciplines must come to terms with the proper way to use social media, without breaching confidentiality or ethical guidelines.

32. The individual receiving services must be given the opportunity to assist in selecting and approving treatment goals. Outcomes must be selected that are primarily aimed at benefiting the individual receiving services.

Advocating for the Client

33. In deciding to provide services, the behavior analyst must determine that services are needed, that medical causes have been ruled out, that the treatment environment will support service delivery, and that a reasonable expectation of success exists.

34. The choice to provide treatment may be divided into two sets of decision rules: (1) determining that the problem is amenable to behavioral intervention and (2) evaluating the likely success of an intervention.

35. The behavior analyst should advocate for the most effective, scientifically supported treatment procedures. The behavior analyst must also take into account the social validity of any proposed treatment procedures. The behavior analyst has the responsibility to objectively evaluate alternative treatments when necessary and appropriate.

36. Behavior analysts embrace the scientific method, and promote evidence-based practices in all consumer-related services.

Conflict of Interest

37. A conflict of interest occurs when a principal party, alone or in connection with family, friends, or associates, has a vested interest in the outcome of the interaction.

38. All sources of conflict of interest, particularly dual relationships, are to be avoided.

Creating a Culture of Ethical Practice

39. When coordinating with other professional service providers, behavior analysts have a responsibility to review and appraise the effects of any treatments that might affect the goals of the behavior change program.

40. To foster a culture of ethical practice, organizations can commit to providing ethics-related trainings, support systems, and differential reinforcement of ethical behaviors.

41. Continuous education and discussion about ethics and the contingencies that support analysis and assistance seeking are core features of an environment designed to maintain ethical behavior in an organization.

KEY TERMS

confidentiality
conflict of interest
ethical codes of behavior
ethics
informed consent

MULTIPLE-CHOICE QUESTIONS

1. When trying to determine the right thing to do in a given situation, which factors should you consider?
   - **a.** Your own personal history in similar circumstances
   - **b.** Your professional training and experience
   - **c.** The rules of the local setting
   - **d.** Ethical guidelines published by your professional organization
   - **e.** All of these

   Hint: (See "What Is Ethics and Why Is It Important?")

2. "Ethics" encompasses:
   - **a.** The primary goal is of furthering the welfare of the client
   - **b.** Behaviors and decisions that revolve around the "right thing to do"
   - **c.** Following the rules of professional codes and standards
   - **d.** What is morally right
   - **e.** All of these

   Hint: (See "What Is Ethics and Why Is It Important?")

3. Which of the following organizations does not have ethical guidelines, standards, or position statements that behavior analysts should follow, according to your text?
   - **a.** American Psychological Association
   - **b.** Association for Behavior Analysis
   - **c.** Association for Direct Instruction
   - **d.** Behavior Analyst Certification Board

   Hint: (See "Standards of Professional Practice for Applied Behavior Analysts")

4. Which of the following is not a right to effective education, as delineated by the Association for Behavior Analysis in 1990?
   - **a.** Social and physical school environments that encourage and maintain academic achievement and progress
   - **b.** Curriculum and objectives based on empirically validated hierarchies or sequences of instructional objectives
   - **c.** Financial consequences for the school and its performance are based on subjective measures by the school board
   - **d.** Specific mastery criteria that include both the accuracy and speed dimensions of fluent performance

   Hint: (See "Standards of Professional Practice for Applied Behavior Analysts")

5. Which of these does not help individuals gain professional competence in their field?
   - **a.** Formal coursework
   - **b.** Supervised practica
   - **c.** Mentored professional experience
   - **d.** Attending nonaccredited workshops and training
   - **e.** Successfully completing accredited university training programs

   Hint: (See "Ensuring Professional Competence")

6. What should you do if you are in a situation where the demands of the situation exceed your training?
   - **a.** Reassure the clients that you are a qualified behavior analyst and continue with the intervention you have in place.
   - **b.** Tell the clients that you are not qualified to provide treatment and walk away.
   - **c.** Seek additional training in the area of competence so as to improve your skills.
   - **d.** Make a referral to another behavior analyst or consultant.
   - **e.** The third and fourth answer choices are correct.

   Hint: (See "Ensuring Professional Competence")

7. Joan is a behavior analyst providing services to a child with autism. She has been having a difficult time with the child's family. She is out for coffee one day, when her cell phone rings. It is her mother, who asks how things are going. She tells her mother about her woes with this family, explaining that the intervention she has implemented doesn't seem to be working and that she is having trouble with the parents, who seem to be questioning her skills. Her mother, who is a retired teacher (not a behavior analyst), asks, "Is this the same family that you had trouble with last year?" Joan replies, "No, this is a different family. This is the Jacksons. They are very wealthy and live on the upper side of town. You remember them. You met them at a school board meeting once." Consider Joan's behavior relative to issues of confidentiality. Which of the following statements is true?
   - **a.** Joan committed a breach of confidentiality unintentionally, so this breech of confidentiality is not of concern.
   - **b.** Joan committed a breach of confidentiality intentionally, and this is of great concern.
   - **c.** Joan committed a breach of confidentiality intentionally, but it is not of great concern because her goal was to serve the greater good and avoid future problems for this child.
   - **d.** No breach of confidentiality occurred because Joan did not give out any personally identifiable information.

   Hint: (See "Ethical Issues in Client Services")

8. Michael is a behavior analyst working at an adult day treatment program, where the staff always involves their clients in planning their day programs. Today, Michael is holding a meeting with one of his clients, Dan, who has moderate mental disabilities and has been having trouble with his behavior during work periods. Dan has been getting up partway through the day and running out of the room, screaming. While he is out of the room, he often tears papers off the wall and destroys the agency's property. The team is meeting with Dan today to determine how his intervention will be revised. Michael is proposing that a response cost intervention be imposed, where Dan will lose some of the money he has earned for every minute

he is out of the room. Dan is opposed to this intervention because he doesn't want to be fined. Michael and the team say to Dan, "But this intervention is trying to help you. I guess if you don't want to do this, we will have to revise your intervention so that you don't earn as much money. We need to pay for all the damage you have caused." After discussing some other options that the team didn't feel would be as effective, Dan says he'd like to think about it, but his bus just pulled up and he needs to leave for the day. The team tells him this needs to be decided today, or he cannot return to the facility tomorrow. What ethical issue has the team violated?

a. They have implemented intervention and assessment without Dan's consent.

b. They are attempting to coerce Dan into the intervention, interfering with his right to voluntarily participate in the intervention.

c. They have breeched confidentiality by sharing this information with Dan and staff.

d. Michael has not shared all of the relevant treatment information with Dan so that he can make an informed decision.

Hint: (See "Ethical Issues in Client Services")

9. Which of these are important considerations in deciding whether or not an intervention is likely to be successful?

a. Is the client willing to participate?

b. Has the behavior been successfully treated in the research literature?

c. Is public support likely?

d. Will those involved in the treatment have control over the critical contingencies?

e. All of these

Hint: (See "Advocating for the Client")

10. What is a conflict of interest?

a. When a behavior analyst, either alone or in tandem with a client, family, friends or associates of the client has a vested interest in the outcome of the interaction.

b. When all members of a team have a similar treatment goal for an individual.

c. When one member of a team has a negative history with the behavior analyst and the two do not get along well.

d. When parents hire a behavior analyst to consult with their child and to provide recommendations to a local school.

Hint: (See "Advocating for the Client")

ESSAY-TYPE QUESTIONS

1. Why are ethics important? Specifically, what do ethical principles do to help guide the practitioner?

Hint: (See "What Is Ethics and Why Is It Important?")

2. Explain the role of professional organizations in creating and enforcing professional ethical codes.

Hint: (See "Standards of Professional Practice for Applied Behavior Analysts")

3. Give an example of individuals practicing within their area of competence and an example of persons practicing outside their area of competence.

Hint: (See "Ensuring Professional Competence")

4. Give an example of a situation where you might provide treatment without consent from either an individual or a guardian.

Hint: (See "Ethical Issues in Client Services")

5. Prior to initiating treatment, the behavior analyst must take into account a number of important considerations. List those considerations.

Hint: (See "Advocating for the Client")

NOTES

1. The Behavior Analyst Certification Board's (BACB's) Professional and Ethical Compliance Code for Behavior Analysts (the "Code") consolidates, updates, and replaces the BACB's Professional Disciplinary and Ethical Standards and Guidelines for Responsible Conduct for Behavior Analysts. *The Compliance Code* includes 10 sections relevant to professional and ethical behavior of behavior analysts, along with a glossary of terms. Effective January 1, 2016, all BACB applicants, certificants, and registrants will be required to adhere to *The Compliance Code*. Source: www.BACB.com.
2. For a list of colleges and universities with graduate programs in behavior analysis, see the *ABAI's Behavior Analysis Training Directory* [https://www.abainternational.org/constituents/educators/training-directory.aspx].
3. Information on the BACB's certification requirements and process is available at: www.BACB.com.
4. Readers interested in early literature on risk–benefit analysis should see Spreat (1982) and Axelrod, Spreat, Berry, & Moyer (1993).

Epilogue

From its earliest days, the story of applied behavior analysis has been one of continuous advancements in understanding how behavior works and how to change it for the better. Today's behavior analysts help more people learn new skills, achieve performance goals, and replace problem behaviors with healthy alternatives in a wider range of settings and circumstances than ever before. We are confident tomorrow's behavior analysts will be even more skillful and effective in improving the quality of people's lives.

Some of you will earn advanced degrees in behavior analysis. Your research, writing, and teaching will expand the knowledge base and practice of the science. Many of you will become professionals whose practice is informed by applied behavior analysis. Your skillful and ethical use of behavioral interventions will enhance the quality of the lives you touch. We hope all of you will bring what you have learned about behavior to your everyday lives.

We cannot express our perspective on the potential of, and necessity for, applied behavior analysis to contribute to a healthier and happier human experience more eloquently than did Ellen Reese (1966) more than 50 years ago:

> The pertinent question is how to use our increasing knowledge of behavior to best advantage. . . . It is a question that concerns all of us, as all of us exercise control over our families and friends, and all of us delegate control to professionals in education, government, and law. We can close our eyes to the techniques of behavior control and risk everything, or we can use these techniques effectively to create the kind of world we want to live in. The latter course requires that we keep informed. Fortunately, the science of behavior is neither esoteric nor obscure. It is a young science, an exciting one, and an important one; for surely a proper study of mankind is his behavior. (p. 63)

Applied behavior analysis is still a young, exciting, and vital science. Given humanity's current challenges, its potential for good is more important than ever before.

Thank you for reading our book. We hope you will continue learning about applied behavior analysis and use your growing knowledge to help make the world a better place. These endeavors are worthy of your talent, time, and effort.

GLOSSARY

A

A-B design A two-phase experimental design consisting of a pretreatment baseline condition (A) followed by a treatment condition (B).

A-B-A design A three-phase experimental design consisting of an initial baseline phase (A) until steady state responding (or counter-therapeutic trend) is obtained, an intervention phase in which the treatment condition (B) is implemented until the behavior has changed and steady state responding is obtained, and a return to baseline conditions (A) by withdrawing the independent variable to see whether responding "reverses" to levels observed in the initial baseline phase. (See also **A-B-A-B design, reversal design,** and **withdrawal design.**)

A-B-A-B design An experimental design consisting of (1) an initial baseline phase (A) until steady state responding (or counter-therapeutic trend) is obtained, (2) an initial intervention phase in which the treatment variable (B) is implemented until the behavior has changed and steady state responding is obtained, (3) a return to baseline conditions (A) by withdrawing the independent variable to see whether responding "reverses" to levels observed in the initial baseline phase, and (4) a second intervention phase (B) to see whether initial treatment effects are replicated. (Also called *reversal design* and *withdrawal design.*)

abative effect (of a motivating operation) A decrease in the current frequency of behavior that has been reinforced by some stimulus, object, or event whose reinforcing effectiveness depends on the same motivating operation. For example, food ingestion abates (decreases the current frequency of) behavior such as opening the fridge that has been reinforced by food.

ABC recording See **anecdotal observation.**

abolishing operation (AO) A motivating operation that decreases the reinforcing effectiveness of a stimulus, object, or event. For example, the reinforcing effectiveness of food is abolished as a result of food ingestion.

acceptance and commitment therapy (ACT) An evidence-based behavior therapy focusing on general well-being, defined as making reliable contact with high-priority positive reinforcers.

accuracy (of measurement) The extent to which observed values, the data produced by measuring an event, match the true state, or true values, of the event as it exists in nature. (See also **observed value** and **true value.**)

adapted alternating treatments design A variation of the multielement design for comparing the efficiency of instructional procedures. The comparison phase of the design features the alternating application of two (usually) or more different teaching methods, each method applied to different but equivalent sets of instructional items. All items are topographically different members of the same response or skill class, such as reading printed words, defining vocabulary terms, spelling words, answering math problems, and stating history facts.

add-in component analysis A method for conducting a component analysis in which components are assessed individually or in combination before the complete treatment package is presented. The add-in method can identify sufficient components. Sequence and floor or ceiling effects may mask the effects of components added in toward the end of the analysis.

adjunctive behavior Behavior that occurs as a collateral effect of a schedule of periodic reinforcement for other behavior; time-filling or interim activities (e.g., doodling, idle talking, smoking, drinking) that are induced by schedules of reinforcement during times when reinforcement is unlikely to be delivered. (Also called *schedule-induced behavior*.)

affirmation of the consequent A three-step form of reasoning that begins with a true antecedent–consequent (if-A-then-B) statement and proceeds as follows: (1) If A is true, then B is true; (2) B is found to be true; (3) therefore, A is true. Although other factors could be responsible for the truthfulness of A, a sound experiment affirms several if-A-then-B possibilities, each one reducing the likelihood of factors other than the independent variable being responsible for the observed changes in behavior.

alternating treatments design See **multielement design.**

alternative schedule (alt) Provides reinforcement when the response requirements of any of two or more simultaneously available component schedules are met.

anecdotal observation A form of direct, continuous observation in which the observer records a descriptive, temporally sequenced account of all behavior(s) of interest and the antecedent conditions and consequences for those behaviors as those events occur in the client's natural environment. (Also called *ABC recording*.)

antecedent An environmental condition or stimulus change existing or occurring prior to a behavior of interest.

antecedent exercise An antecedent intervention, implemented independently of occurrences of the problem behavior, that usually has clients engage in some effortful form of aerobic activity (e.g., walking, jogging, dancing, calisthenics, roller skating). Applied behavior analysts have used antecedent exercise in the treatment of many maladaptive behaviors such as self-injurious behavior (SIB), aggression, and diverse behaviors such as inappropriate vocalizations, repetitive movements, talking-out, out-of-seat, and stereotypic behaviors.

antecedent intervention A behavior change strategy that manipulates antecedent stimuli based on (a) motivating operations (evocative and abative effects), (b) stimulus control (differential availability of reinforcement), and (c) contingency-independent interventions (e.g., protective equipment, and restraint). (See also **functional communication training, high-probability request sequence,** and **noncontingent reinforcement.** Contrast with **antecedent control,** a behavior change intervention that manipulates contingency-dependent consequence events to affect stimulus control.)

antecedent stimulus class A set of stimuli that share a common relationship. All stimuli in an antecedent stimulus class evoke the same operant behavior, or elicit the same respondent behavior. (See also **arbitrary stimulus class** and **feature stimulus class.**)

applied behavior analysis (ABA) The science in which tactics derived from the principles of behavior are applied to improve socially significant behavior and experimentation is used to identify the variables responsible for the improvement in behavior.

arbitrarily applicable relational responding (AARR) Forming new stimulus classes with little or no reinforced practice.

arbitrary relations Stimuli to which people respond in interlocked ways, not because of physical similarity, but because social-verbal reinforcement contingencies teach people to respond to them in this way.

arbitrary stimulus class Antecedent stimuli that evoke the same response but do not resemble each other in physical form or share a relational aspect such as bigger or under (e.g., peanuts, cheese, coconut milk, and chicken breasts are members of an arbitrary stimulus class if they evoke the response "sources of protein"). (Compare to **feature stimulus class.**)

artifact An outcome or result that appears to exist because of the way it is measured but in fact does not correspond to what actually occurred.

ascending baseline A data path that shows an increasing trend in the response measure over time. (Compare to **descending baseline.**)

audience Anyone who functions as a discriminative stimulus evoking verbal behavior. Different audiences may control different verbal behavior about the same topic because of a differential reinforcement history. Teens may describe the same event in different ways when talking to peers versus parents.

autoclitic The autoclitic relation involves two interlocking levels of verbal behavior emitted in one utterance. One level is a primary response (e.g., "The ice is solid"), while the other type is the secondary autoclitic response (e.g., adding "I think"). Autoclitic behavior benefits the listener by providing additional information regarding the primary response.

autoclitic frame Provide structure among verbal operants in terms of order, agreement, grouping, and composition of larger units of verbal behavior such as sentences (see Palmer, 2016). Autoclitic frames help speakers generate novel

utterances. For example, experience with the frame "the girl's (ball, dog, coat) may enable a child to say "the girl's hat" without prior teaching.

autoclitic mand An autoclitic mand involves supplemental control by an MO related to some aspect of the speaker's primary verbal response. The autoclitic mand (e.g., "Believe me, they are wrong.") enjoins a listener to take some specific action regarding the primary response, and the autoclitic behavior is reinforced by that action.

autoclitic tact An autoclitic tact is controlled by some nonverbal feature of the primary response or its controlling variables, and the autoclitic response informs the listener of that feature. This behavior is reinforced by listeners who provide generalized reinforcement.

automatic contingencies Skinner (1957) used "automatic" to identify circumstances in which behavior is evoked, shaped, maintained, or weakened by environmental variables occurring without direct manipulation by other people. All behavioral principles (e.g., reinforcement, extinction, punishment) can affect our behavior automatically.

automatic punishment Punishment that occurs independent of the social mediation by others (i.e., a response product serves as a punisher independent of the social environment).

automatic reinforcement Reinforcement that occurs independent of the social mediation of others (e.g., scratching an insect bite relieves the itch).

automaticity (of reinforcement) Refers to the fact that behavior is modified by its consequences irrespective of the person's awareness; a person does not have to recognize or verbalize the relation between her behavior and a reinforcing consequence, or even know that a consequence has occurred, for reinforcement to "work." (Contrast with **automatic reinforcement.**)

aversive stimulus In general, an unpleasant or noxious stimulus; more technically, a stimulus change or condition that functions (a) to evoke a behavior that has terminated it in the past, (b) as a punisher when presented following behavior, and/or (c) as a reinforcer when withdrawn following behavior.

avoidance contingency A contingency in which a response prevents or postpones the presentation of a stimulus. (Compare to **escape contingency.**)

B

B-A-B design A three-phase experimental design that begins with the treatment condition. After steady state responding has been obtained during the initial treatment phase (B), the treatment variable is withdrawn (A) to see whether responding changes in the absence of the independent variable. The treatment variable is then reintroduced (B) in an attempt to recapture the level of responding obtained during the first treatment phase.

backup reinforcers Preferred items, activities, or privileges that participants obtain by exchanging earned tokens in a token economy.

backward chaining A teaching procedure in which a trainer completes all but the last behavior in a chain, which is performed by the learner, who then receives reinforcement for completing the chain. When the learner shows competence in performing the final step in the chain, the trainer performs all but the last two behaviors in the chain, the learner emits the final two steps to complete the chain, and reinforcement is delivered. This sequence is continued until the learner completes the entire chain independently.

backward chaining with leaps ahead A backward chaining procedure in which some steps in the task analysis are skipped; used to increase the efficiency of teaching long behavior chains when there is evidence that the skipped steps are in the learner's repertoire.

bar graph A simple and versatile graphic format for summarizing behavioral data; shares most of the line graph's features except that it does not have distinct data points representing successive response measures through time. (Also called a *histogram*.)

baseline A condition of an experiment in which the independent variable is not present; data obtained during baseline are the basis for determining the effects of the independent variable; a control condition that does not necessarily mean the absence of instruction or treatment, only the absence of a specific independent variable of experimental interest.

baseline logic A term sometimes used to refer to the experimental reasoning inherent in single-subject experimental designs; entails three elements: prediction, verification, and replication. (See also **steady state strategy.**)

behavior That portion of an organism's interaction with its environment that involves movement of some part of the organism (Johnston & Pennypacker, 2009, p. 31). (See also **operant behavior, respondent behavior, response,** and **response class.**)

Behavior Analyst Certification Board's (BACB's) Professional and Ethical Compliance Code for Behavior Analysts Consolidates, updates, and replaces the BACB's Professional Disciplinary and Ethical Standards and Guidelines for Responsible Conduct for Behavior Analysts. *The Compliance Code* includes 10 sections relevant to professional and ethical behavior of behavior analysts, along with a glossary of terms. Effective January 1, 2016, all BACB applicants, certificants, and registrants will be required to adhere to *The Compliance Code*.

behavior chain A sequence of responses in which each response produces a stimulus change that functions as conditioned reinforcement for that response and as a discriminative stimulus for the next response in the chain; reinforcement for the last response in a chain maintains the reinforcing effectiveness of the stimulus changes produced by all previous responses in the chain.

behavior chain interruption strategy An intervention that relies on the participant's skill to initially perform all the critical elements of a chain independently, but the chain is then interrupted, or a link in the chain is made unavailable at a predetermined time, so that another behavior can be prompted or emitted.

behavior chain with a limited hold A contingency that specifies a time interval by which a behavior chain must be completed for reinforcement to be delivered.

behavior change tactic A technologically consistent method for changing behavior derived from one or more principles of behavior (e.g., response cost is derived from the principle of negative punishment); possesses sufficient generality across subjects, settings, and/or behaviors to warrant its codification and dissemination.

behavior checklist An itemized list that provides descriptions of specific skills (usually in hierarchical order) and the conditions under which each skill should be observed. Some checklists are designed to assess one particular behavior or skill area. Others address multiple behaviors or skill areas. Most use a Likert scale to rate responses.

behavior trap An interrelated community of contingencies of reinforcement that can be especially powerful, producing substantial and long-lasting behavior changes. Effective behavior traps share four essential features: (a) They are "baited" with virtually irresistible reinforcers that "lure" the student to the trap; (b) only a low-effort response already in the student's repertoire is necessary to enter the trap; (c) once inside the trap, interrelated contingencies of reinforcement motivate the student to acquire, extend, and maintain targeted academic and/or social skills; and (d) they can remain effective for a long time because students show few, if any, satiation effects.

behavior-altering effect (of a motivating operation) Either (a) an increase in the current frequency of behavior that has been reinforced by some stimulus, object, or event, called an evocative effect; or (b) a decrease in the current frequency of behavior that has been reinforced by some stimulus, object, or event, called an abative effect. For example, the current frequency of behavior that has been reinforced with food, such as opening the fridge, is evoked (increased) or abated (decreased) by food deprivation or food ingestion, respectively.

behavioral assessment A form of assessment that includes indirect and direct procedures such as interviews, checklists, and tests to identify and define the specific target behavior. In addition to identifying behavior(s) to change, comprehensive behavioral assessment can uncover functional relations between variables; it provides context on the resources, assets, significant others, competing contingencies, maintenance and generalization factors, and reinforcers (or punishers) that can be combined to improve the efficiency of an intervention.

behavioral contract See **contingency contract.**

behavioral contrast The phenomenon in which a change in one component of a multiple schedule that increases or decreases the rate of responding on that

component is accompanied by a change in the response rate in the opposite direction on the other, unaltered component of the schedule.

behavioral cusp A behavior that has sudden and dramatic consequences that extend well beyond the idiosyncratic change itself because it exposes the person to new environments, reinforcers, contingencies, responses, and stimulus controls. (See also **pivotal behavior.**)

behavioral inflexibility An insensitivity to external stimuli occurring when private events interfere with well-being behaviors on which high-priority positive reinforcers are contingent.

behavioral momentum Describes the resistance to change in a behavior's rate of responding following an alteration in reinforcement conditions. The momentum metaphor has also been used to describe the effects produced by the **high-probability (high-*p*) request sequence.**

behaviorism The philosophy of a science of behavior; there are various forms of behaviorism. (See also **methodological behaviorism** and **radical behaviorism.**)

believability The extent to which the researcher convinces herself and others that the data are trustworthy and deserve interpretation. Measures of interobserver agreement (IOA) are the most often used index of believability in applied behavior analysis. (See also **interobserver agreement [IOA].**)

bidirectional naming (BiN) A higher-order verbal cusp consisting of the fusing together of the speaker and listener repertoires in bidirectional relations (Horne & Lowe, 1996). A new word acquired as listener can generate a tact without further training, and a new word acquired as a tact can generate a listener relation without further training (these effects are consistent with emergent symmetry and mutual entailment).

bonus response cost A procedure for implementing response cost in which the person is provided a reservoir of reinforcers that are removed in predetermined amounts contingent on the occurrence of the target behavior.

brief functional analysis An analysis in which only one or two 5- to 10-min sessions are conducted for each condition. A convincing demonstration of function may be achieved by either alternating a condition that produces problem behavior with one that does not or conducting a contingency reversal. Brief functional analysis may reveal a functional relation in fewer sessions than a full functional analysis.

C

calibration The process of comparing the data produced by a measurement system to a known standard or true value and, when sources of error are found, using that information to correct or improve the measurement system.

causal relations If-then relationships (e.g., if A, then B; if B, then C) that are a central feature of understanding and doing science. With respect to stimulus relations, causal relations can define the structure of a stimulus class or define the behavior function through which stimuli in a class are transformed.

celeration The change (acceleration or deceleration) in rate of responding over time; based on count per unit of time (rate); expressed as a factor by which responding is accelerating or decelerating (multiplying or dividing); displayed with a trend line on a Standard Celeration Chart. *Celeration* is a generic term without specific reference to accelerating or decelerating rates of response. (See also **Standard Celeration Chart.**)

celeration time period A unit of time (e.g., per week, per month) in which celeration is plotted on a Standard Celeration Chart. (See also **celeration** and **celeration trend line.**)

celeration trend line The celeration trend line is measured as a factor by which rate multiplies or divides across the celeration time periods (e.g., rate per week, rate per month, rate per year, and rate per decade). (See also **celeration.**)

chained schedule A schedule of reinforcement in which the response requirements of two or more basic schedules must be met in a specific sequence before reinforcement is delivered; a discriminative stimulus is correlated with each component of the schedule.

chaining Various methods for linking specific sequences of stimuli and responses to form new performances. (See also **backward chaining, backward chaining with leaps ahead, behavior chain**, and **forward chaining.**)

changing criterion design An experimental design in which an initial baseline phase is followed by a series of treatment phases consisting of successive and gradually changing criteria for reinforcement or punishment. Experimental control is evidenced by the extent the level of responding changes to conform to each new criterion.

class expansion A new member is added to a demonstrated stimulus equivalence class as the result of teaching a new conditional discrimination.

class merger Independent equivalence classes are combined as the result of teaching a new but interrelated conditional discrimination.

class-specific reinforcement A match-to-sample procedure in which not only is the correct comparison choice conditional on the sample stimulus, but the type of consequence delivered is, too; class-specific consequences themselves become members of the equivalence classes.

clicker training A term popularized by Pryor (1999) for shaping behavior using conditioned reinforcement in the form of an auditory stimulus. A handheld device produces a click sound when pressed. The trainer pairs other forms of reinforcement (e.g., edible treats) with the click sound so that the sound becomes a conditioned reinforcer.

codic A type of verbal behavior where the form of the response is under the functional control of a verbal stimulus with point-to-point correspondence, but without formal similarity. There is also a history of generalized reinforcement.

combinatorial entailment A relation involving two stimuli that both participate in mutual entailment with some common third stimulus (See also **mutual entailment**).

Compliance Code See the **Behavior Analyst Certification Board's (BACB's) Professional and Ethical Compliance Code for Behavior Analysts.**

component analysis Any experiment designed to identify the active elements of a treatment condition, the relative contributions of different variables in a treatment package, and/or the necessary and sufficient components of an intervention. Component analyses take many forms, but the basic strategy is to compare levels of responding across successive phases in which the intervention is implemented with one or more components left out.

compound schedule A schedule of reinforcement consisting of two or more elements of continuous reinforcement (CRF), the four intermittent schedules of reinforcement (FR, VR, FI, VI), differential reinforcement of various rates of responding (DRH, DRL), and extinction. The elements from these basic schedules can occur successively or simultaneously and with or without discriminative stimuli; reinforcement may be contingent on meeting the requirements of each element of the schedule independently or in combination with all elements.

compound verbal discrimination Involves two or more verbal S^Ds (convergent multiple control) that each independently evoke behavior, but when they both occur in the same antecedent configuration, a different S^D is generated, and a more specific behavior is evoked.

concept A stimulus class whose members share a common set of features.

concept formation A complex example of stimulus control that requires stimulus generalization within a class of stimuli and discrimination between classes of stimuli.

concurrent chains design An experimental design in which participants are presented with two or more response options; each option is associated with a distinctive discriminative stimulus and leads to a different set of treatment procedures. (Also called *concurrent schedule design.*)

concurrent schedule (conc) A schedule of reinforcement in which two or more contingencies of reinforcement (elements) operate independently and simultaneously for two or more behaviors.

conditional discrimination Performance in a match-to-sample procedure in which discrimination between the comparison stimuli is conditional on, or depends on, the sample stimulus present on each trial.

conditional probability The likelihood that a target behavior will occur in a given circumstance; computed by calculating (a) the proportion of occurrences of behavior that were preceded by a specific antecedent variable and (b) the proportion of occurrences of problem behavior that were followed by a specific consequence. Conditional probabilities range from 0.0 to 1.0; the closer the conditional probability is to 1.0, the stronger the relationship is between the target behavior and the antecedent/consequence variables.

conditioned motivating operation (CMO) A motivating operation whose value-altering effect depends on a learning history. For example, because of the relation between locked doors and keys, having to open a locked door is a CMO that makes keys more effective as reinforcers, and evokes behavior that has obtained such keys.

conditioned negative reinforcer A previously neutral stimulus change that functions as a negative reinforcer because of prior pairing with one or more negative reinforcers. (See also **negative reinforcer**; compare to **unconditioned negative reinforcer.**)

conditioned punisher A previously neutral stimulus change that functions as a punisher because of prior pairing with one or more other punishers. (Sometimes called *secondary* or *learned punisher*; compare to **unconditioned punisher.**)

conditioned reflex A learned stimulus–response functional relation consisting of an antecedent stimulus (e.g., sound of refrigerator door opening) and the response it elicits (e.g., salivation); each person's repertoire of conditioned reflexes is the product of his or her history of interactions with the environment (ontogeny). (See also **respondent conditioning** and **unconditioned reflex.**)

conditioned reinforcer A stimulus change that functions as a reinforcer because of prior pairing with one or more other reinforcers. (Sometimes called *secondary* or *learned reinforcer*.)

conditioned stimulus (CS) The stimulus component of a conditioned reflex; a formerly neutral stimulus change that elicits respondent behavior only after it has been paired with an unconditioned stimulus (US) or another CS.

confidentiality Describes a situation of trust insofar as any information regarding a person receiving or having received services may not be discussed with or otherwise made available to another person or group, unless that person has provided explicit authorization, usually written and signed, for release of such information.

conflict of interest A situation in which a person in a position of responsibility or trust has competing professional or personal interests that make it difficult to fulfill his or her duties impartially.

confounding variable An uncontrolled factor known or suspected to exert influence on the dependent variable.

conjunctive schedule (conj) A schedule of reinforcement that is in effect whenever reinforcement follows the completion of response requirements for two or more schedules of reinforcement.

consequence A stimulus change that follows a behavior of interest. Some consequences, especially those that are immediate and relevant to current motivational states, have significant influence on future behavior; others have little effect. (See also **punisher** and **reinforcer.**)

constant time delay A procedure for transferring stimulus control from contrived response prompts to naturally existing stimuli. After the student has responded correctly to several 0-sec delay trials, after which presentation of the response prompt follows the instructional stimulus by a predetermined and fixed delay (usually 3 or 4 seconds) for all subsequent trials. (See also **progressive time delay**).

contextual control The situation or context in which a stimulus (or stimulus class) occurs determines its function. More specifically: a type of stimulus control requiring three levels of antecedent stimuli, such that the functions of the stimuli in a conditional discrimination vary depending on the context. Contextual control training requires a five-term contingency. It allows for the same stimuli to be members of more than one equivalence class, depending on the context.

contextual stimulus Signals the type of relational responding that will be reinforced.

contingency Refers to dependent and/or temporal relations between operant behavior and its controlling variables. (See also **contingent** and **three-term contingency.**)

contingency contract A mutually agreed-upon document between parties (e.g., parent and child) describing a contingent relationship between the completion of specified behavior(s) and access to specified reinforcer(s). (Also called **behavioral contract.**)

contingency reversal Exchanging the reinforcement contingencies for two topographically different responses. For example, if Behavior A results in reinforcement on an FR 1 schedule of reinforcement and Behavior B results in reinforcement being withheld (extinction), a contingency reversal consists of changing the contingencies such that Behavior A now results in extinction and Behavior B results in reinforcement on an FR 1 schedule.

contingency-shaped behavior Behavior acquired by direct experience with contingencies.

contingent Describes reinforcement (or punishment) that is delivered only after the target behavior has occurred.

contingency space analysis A graphic display of the probability of one event (e.g., a particular consequence) given the occurrence (or not) of another event (e.g., the presence or absence of a particular behavior). Contingencies are considered positive (i.e., a specific consequence is more likely to occur), negative (i.e., a specific consequence is less likely to occur), or neutral (i.e., a specific consequence is neither more or less likely to occur).

contingent observation A procedure for implementing time-out in which the person is repositioned within an existing setting such that observation of ongoing activities remains, but access to reinforcement is lost.

continuous measurement Measurement conducted in a manner such that all instances of the response class(es) of interest are detected during the observation period.

continuous reinforcement (CRF) A schedule of reinforcement that provides reinforcement for each occurrence of the target behavior.

contrived contingency Any contingency of reinforcement (or punishment) designed and implemented by a behavior analyst or practitioner to achieve the acquisition, maintenance, and/or generalization of a targeted behavior change. (Contrast with **naturally existing contingency.**)

contrived mediating stimulus Any stimulus made functional for the target behavior in the instructional setting that later prompts or aids the learner in performing the target behavior in a generalization setting.

copying text An elementary verbal operant involving a written response that is evoked by a written verbal discriminative stimulus that has formal similarity and a history of generalized reinforcement.

count A simple tally of the number of occurrences of a behavior. The observation period, or counting time, should always be noted when reporting *count measures*.

countercontrol Behavior evoked by coercion or aversive forms of external control; takes many forms, including escape, attack, passive resistance; "an emotional reaction of anger or frustration including operant behavior which injures or is otherwise aversive to the controller" (Skinner, 1953, p. 321).

counting time The period of time in which a count of the number of responses emitted is recorded.

cumulative record A type of graph on which the cumulative number of responses emitted is represented on the vertical axis; the steeper the slope of the data path, the greater the response rate.

cumulative recorder A device that automatically draws cumulative records (graphs) that show the rate of response in real time; each time a response is emitted, a pen moves upward across paper that continuously moves at a constant speed.

D

data The results of measurement, usually in quantifiable form; in applied behavior analysis, it refers to measures of some quantifiable dimension of a behavior. The word *data* is the plural form of datum.

data path The level and trend of behavior between successive data points; created by drawing a straight line from the center of each data point in a given data set to the center of the next data point in the same set.

deictic relation A relation between the self, as one stimulus, and other stimuli from the external world.

delay discounting A phenomenon in which delayed rewards, regardless of their significance and magnitude (e.g., enough money for a secure retirement), exert decreasing influence over choice-making behavior as a function of their temporal distance from present circumstances. Both humans and nonhuman

laboratory animals discount the value of delayed rewards; the greater the delay to the reward, the greater the discount (i.e., the less value or influence the reward has on current behavior). (Also called *temporal discounting*.)

delayed multiple baseline design A variation of the multiple baseline design in which an initial baseline, and perhaps intervention, begin for one behavior (or setting, or subject), and subsequent baselines for additional behaviors begin in a staggered or delayed fashion.

dependent group contingency A contingency in which reinforcement for all members of a group is dependent on the behavior of one member of the group or the behavior of a select group of members within the larger group.

dependent variable The measured behavior in an experiment to determine if it changes as a result of manipulations of the independent variable; in applied behavior analysis, it represents some measure of a socially significant behavior. (See also **target behavior;** compare to **independent variable.**)

deprivation The state of an organism with respect to how much time has elapsed since it has consumed or contacted a particular type of reinforcer; also refers to a procedure for increasing the effectiveness of a reinforcer (e.g., withholding a person's access to a reinforcer for a specified period prior to a session). (See also **motivating operation [MO]**; contrast with **satiation.**)

derived stimulus relations Responding indicating a relation (e.g., same as, opposite, different from, better than) between two or more stimuli that emerges as an indirect function of related instruction or experience. (Also called *emergent stimulus relations*.)

descending baseline A data path that shows a decreasing trend in the response measure over time. (Compare to **ascending baseline.**)

descriptive functional behavior assessment Direct observation of problem behavior and the antecedent and consequent events under naturally occurring conditions.

determinism The assumption that the universe is a lawful and orderly place in which phenomena occur in relation to other events and not in a willy-nilly, accidental fashion.

differential reinforcement Reinforcing only those responses within a response class that meet a specific criterion along some dimension(s) (i.e., frequency, topography, duration, latency, or magnitude) and placing all other responses in the class on extinction. (See also **differential reinforcement of alternative behavior**, **differential reinforcement of incompatible behavior**, **differential reinforcement of other behavior**, **stimulus discrimination training**, and **shaping**.)

differential reinforcement of alternative behavior (DRA) A procedure for decreasing problem behavior in which reinforcement is delivered for a behavior that serves as a desirable alternative to the behavior targeted for reduction and withheld following instances of the problem behavior (e.g., reinforcing completion of academic worksheet items when the behavior targeted for reduction is talk-outs).

differential reinforcement of diminishing rates (DRD) A schedule of reinforcement in which reinforcement is provided at the end of a predetermined interval contingent on the number of responses emitted during the interval being fewer than a gradually decreasing criterion based on the individual's performance in previous intervals (e.g., fewer than five responses per 5 minutes, fewer than four responses per 5 minutes, fewer than three responses per 5 minutes).

differential reinforcement of high rates (DRH) A schedule of reinforcement in which reinforcement is provided at the end of a predetermined interval contingent on the number of responses emitted during the interval being greater than a gradually increasing criterion based on the individual's performance in previous intervals (e.g., more than three responses per 5 minutes, more than five responses per 5 minutes, more than eight responses per 5 minutes).

differential reinforcement of incompatible behavior (DRI) A procedure for decreasing problem behavior in which reinforcement is delivered for a behavior that is topographically incompatible with the behavior targeted for reduction and withheld following instances of the problem behavior (e.g., sitting in seat is incompatible with walking around the room).

differential reinforcement of low rates (DRL) A schedule of reinforcement in which reinforcement (a) follows each occurrence of the target behavior that is separated from the previous response by a minimum interresponse time (IRT), or (b) is contingent on the number of responses within a period of time not exceeding a predetermined criterion. Practitioners use DRL schedules to decrease the rate of behaviors that occur too frequently but should be maintained in the learner's repertoire. (See also **full-session DRL, interval DRL,** and **spaced-responding DRL.**)

differential reinforcement of other behavior (DRO) A procedure for decreasing problem behavior in which reinforcement is contingent on the absence of the problem behavior during or at specific times (i.e., momentary DRO; sometimes called *differential reinforcement of zero rates of responding* or *omission training*; see also **fixed-interval DRO, fixed-momentary DRO, variable-interval DRO**, and **variable-momentary DRO.**)

direct measurement Occurs when the behavior that is measured is the same as the behavior that is the focus of the investigation. (Contrast with **indirect measurement.**)

direct replication An experiment in which the researcher attempts to duplicate exactly the conditions of an earlier experiment.

disciplinary standards Statements describing corrective, disciplinary, and revocation actions, depending on the circumstances for violations of a professional; see the BACB Compliance Code.

discontinuous measurement Measurement conducted in a manner such that some instances of the response class(es) of interest may not be detected.

discrete trial Any operant whose response rate is controlled by a given opportunity to emit the response. Each discrete response occurs when an opportunity to respond exists. *Discrete trial, restricted operant*, and *controlled operant* are synonymous technical terms. (Contrast with **free operant.**)

discriminated avoidance A contingency in which responding in the presence of a signal prevents the onset of a stimulus from which escape is a reinforcer. (See also **discriminative stimulus, discriminated operant, free-operant avoidance,** and **stimulus control.**)

discriminated operant An operant that occurs more frequently under some antecedent conditions than under others. (See also **discriminative stimulus [S^D]** and **stimulus control.**)

discriminative stimulus (S^D) A stimulus in the presence of which a given behavior has been reinforced and in the absence of which that behavior has not been reinforced; as a result of this history, an S^D evokes operant behavior because its presence signals the availability of reinforcement (See also **differential reinforcement, stimulus control, stimulus delta [S^Δ],** and **stimulus discrimination training.**)

discriminative stimulus for punishment (S^{Dp}) A stimulus in the presence of which a given behavior has been punished and in the absence of which that behavior has not been punished; as a result of this history, the behavior occurs less often in the presence of the S^{Dp} than in its absence.

distinction relation Responding jointly to two stimuli on the basis of their differences.

double-blind control A procedure that prevents the subject and the experimenter(s) from detecting the presence or absence of the treatment variable; used to eliminate confounding of results by subject expectations, parent and teacher expectations, differential treatment by others, and observer bias. (See also **placebo control.**)

DRI/DRA reversal technique An experimental technique that demonstrates the effects of reinforcement; it uses differential reinforcement of an incompatible or alternative behavior (DRI/DRA) as a control condition instead of a no-reinforcement (baseline) condition. During the DRI/DRA condition, the stimulus change used as reinforcement in the reinforcement condition is presented contingent on occurrences of a specified behavior that is either incompatible with the target behavior or an alternative to the target behavior. A higher level of responding during the reinforcement condition than during the DRI/DRA condition demonstrates that the changes in behavior are the result of *contingent reinforcement*, not simply the presentation of or contact with the stimulus event. (Compare to **DRO reversal technique** and **noncontingent reinforcement [NCR] reversal technique.**)

DRO reversal technique An experimental technique for demonstrating the effects of reinforcement by using differential reinforcement of other behavior (DRO) as a control condition instead of a no-reinforcement (baseline)

condition. During the DRO condition, the stimulus change used as reinforcement in the reinforcement condition is presented contingent on the absence of the target behavior for a specified time period. A higher level of responding during the reinforcement condition than during the DRO condition demonstrates that the changes in behavior are the result of *contingent reinforcement*, not simply the presentation of or contact with the stimulus event. (Compare to **DRI/DRA reversal technique** and **noncontingent reinforcement [NCR] reversal technique.**)

drop-out component analysis A method for conducting a component analysis in which the investigator presents the treatment package and then systematically removes components. If the treatment's effectiveness wanes when a component is removed, the researcher has identified a necessary component. (Contrast with **add-in component analysis**).

duplic A type of verbal behavior where the form of the response is under the functional control of a verbal stimulus with formal similarity, and a history of generalized reinforcement.

duration The total time that behavior occurs; measured by elapsed time from the onset of a response to its end point.

E

echoic An elementary verbal operant involving a vocal response that is evoked by a vocal verbal S^D that has formal similarity between an auditory verbal stimulus and an auditory verbal response product, and a history of generalized reinforcement.

ecological assessment An assessment protocol that acknowledges complex interrelationships between environment and behavior. An ecological assessment is a method for obtaining data across multiple settings and persons.

elementary verbal operants Michael's (1982) term for Skinner's (1957) taxonomy of five different types of speaker behavior (i.e., expressive language) distinguished by their antecedent controlling variables and related history of consequences: mand, tact, intraverbal, duplic, and codic. (See also **echoic**, **copying text, taking dictation**, and **textual behavior**).

emergent stimulus relations Stimulus relations that are not taught directly but emerge as an indirect function of related instruction or experience. (Also called *derived stimulus relations*.)

empiricism The objective observation of the phenomena of interest; objective observations are "independent of the individual prejudices, tastes, and private opinions of the scientist. . . . Results of empirical methods are objective in that they are open to anyone's observation and do not depend on the subjective belief of the individual scientist" (Zuriff, 1985, p. 9).

enriched environment An intervention that provides noncontingent access to preferred sources of reinforcement (e.g., toys, games, social and recreation activities). This noncontingent access to preferred reinforcers arranges a competition between the enriched environment and the stimulation provided by the problem behavior.

environment The conglomerate of real circumstances in which the organism or referenced part of the organism exists; behavior cannot occur in the absence of environment.

environmental enrichment (EE) An antecedent intervention that provides noncontingent access to preferred sources of reinforcement (e.g., toys, games, social and recreation activities). This noncontingent access to preferred sources of reinforcement arranges a competition between the EE intervention and the stimulation provided by the problem behavior

equal-interval scale A scale in which equal distances on the axis represent equal absolute amounts of the variable plotted on the axis (e.g., behavior change on an equal-interval vertical axis).

equivalence test A probe for the emergence of untrained stimulus–stimulus relations that evaluates both symmetry and transitivity simultaneously.

equivalence-class formation The emergence of accurate responding to untrained and nonreinforced stimulus–stimulus relations following the reinforcement of responses to some stimulus–stimulus relations. Requires successful performances on three types of probe trials—reflexivity, symmetry, and transitivity—in the absence of reinforcement. (Sometimes called *stimulus equivalence*.)

errorless learning A variety of techniques for gradually transferring stimulus control with a minimum of errors.

escape contingency A contingency in which a response terminates (produces escape from) an ongoing stimulus. (Compare to **avoidance contingency.**)

escape extinction Behaviors maintained with negative reinforcement are placed on escape extinction when those behaviors are not followed by termination of the aversive stimulus; emitting the target behavior does not enable the person to escape the aversive situation.

establishing operation (EO) A motivating operation that increases the effectiveness of some stimulus, object, or event as a reinforcer. For example, food deprivation establishes food as an effective reinforcer.

ethical codes of behavior Documents generated or adopted by professional organizations that provide clear guidelines for their members when deciding a course of proper action in conducting their professional duties.

ethics Behaviors, practices, and decisions that address such basic and fundamental questions as the following: What is the right thing to do? What's worth doing? What does it mean to be a good behavior analytic practitioner? The *Behavior Analyst Certification Board's (BACB's) Professional and Ethical Compliance Code for Behavior Analysts* details ethical practice for behavior analysts.

event recording Measurement procedure for obtaining a tally or count of the number of times a behavior occurs.

evocative effect (of a motivating operation) An increase in the current frequency of behavior that has been reinforced by some stimulus, object, or event whose reinforcing effectiveness depends on the same motivating operation. For example, food deprivation evokes (increases the current frequency of) behavior such as opening the fridge that has been reinforced by food.

exact count-per-interval IOA The percentage of total intervals in which two observers recorded the same count; the most stringent description of IOA for most data sets obtained by event recording.

exclusion (training) A procedure for building new arbitrary conditional discriminations based on the robust finding that learners will select a novel comparison stimulus over a known one in the presence of a novel sample.

exclusion time-out A procedure for implementing time-out in which, contingent on the occurrence of a target behavior, the person is removed physically from the current environment for a specified period.

experiment A carefully controlled comparison of some measure of the phenomenon of interest (the dependent variable) under two or more different conditions in which only one factor at a time (the independent variable) differs from one condition to another.

experimental analysis of behavior (EAB) A natural science approach to the study of behavior as a subject matter in its own right founded by B. F. Skinner; methodological features include rate of response as a basic dependent variable, repeated or continuous measurement of clearly defined response classes, within-subject experimental comparisons instead of group design, visual analysis of graphed data instead of statistical inference, and an emphasis on describing functional relations between behavior and controlling variables in the environment over formal theory testing.

experimental control Two meanings: (a) the outcome of an experiment that demonstrates convincingly a functional relation, meaning that experimental control is achieved when a predictable change in behavior (the dependent variable) can be reliably produced by manipulating a specific aspect of the environment (the independent variable); and (b) the extent to which a researcher maintains precise control of the independent variable by presenting it, withdrawing it, and/or varying its value, and also by eliminating or holding constant all confounding and extraneous variables. (See also **confounding variable, extraneous variable,** and **independent variable.**)

experimental design The particular type and sequence of conditions in a study so that meaningful comparisons of the effects of the presence and absence (or different values) of the independent variable can be made.

explanatory fiction A fictitious or hypothetical variable that often takes the form of another name for the observed phenomenon it claims to explain and contributes nothing to a functional account or understanding of the phenomenon, such as "intelligence" or "cognitive awareness" as explanations for why

an organism pushes the lever when the light is on and food is available but does not push the lever when the light is off and no food is available.

external validity The degree to which a study's findings have generality to other subjects, settings, and/or behaviors. (Compare to **internal validity.**)

extinction The discontinuing of a reinforcement of a previously reinforced behavior (i.e., responses no longer produce reinforcement); the primary effect is a decrease in the frequency of the behavior until it reaches a prereinforced level or ultimately ceases to occur. (See also **extinction burst** and **spontaneous recovery;** compare to **respondent extinction.**)

extinction burst An increase in the frequency of responding when an extinction procedure is initially implemented.

extinction-induced variability Phenomenon in which diverse and novel forms of behavior are sometimes observed during the extinction process.

extraneous variable Any aspect of the experimental setting (e.g., lighting, temperature) that must be held constant to prevent unplanned environmental variation.

F

fading A procedure for transferring stimulus control in which features of an antecedent stimulus (e.g., shape, size, position, color) controlling a behavior are gradually changed to a new stimulus while maintaining the current behavior; stimulus features can be faded in (enhanced) or faded out (reduced).

feature stimulus class Stimuli that share common physical forms or structures (e.g., made from wood, four legs, round, blue) or common relative relationships (e.g., bigger than, hotter than, higher than, next to). (Compare to **arbitrary stimulus class.**)

fixed interval (FI) A schedule of reinforcement in which reinforcement is delivered for the first response emitted following the passage of a fixed duration of time since the last response was reinforced (e.g., on an FI 3-min schedule, the first response following the passage of 3 minutes is reinforced).

fixed ratio (FR) A schedule of reinforcement requiring a fixed number of responses for reinforcement (e.g., an FR 4 schedule of reinforcement follows every fourth response).

fixed-interval DRO (FI-DRO) A DRO procedure in which reinforcement is available at the end of intervals of fixed duration and delivered contingent on the absence of the problem behavior during each interval. (See also **differential reinforcement of other behavior [DRO].**)

fixed-momentary DRO (FM-DRO) A DRO procedure in which reinforcement is available at specific moments in time, which are separated by a fixed amount of time, and delivered contingent on the problem not occurring at those moments. (See also **differential reinforcement of other behavior [DRO].**)

fixed-time schedule (FT) A schedule for the delivery of noncontingent stimuli in which a time interval remains the same from one delivery to the next.

formal similarity Occurs when the controlling antecedent stimulus and the response or response product (a) share the same sense mode (e.g., both stimulus and response are visual, auditory, or tactile) and (b) physically resemble each other. Verbal relations with formal similarity are echoic, copying a text, and imitation as it relates to sign language.

forward chaining A method for teaching behavior chains that begins with the learner being prompted and taught to perform the first behavior in the task analysis; the trainer completes the remaining steps in the chain. When the learner shows competence in performing the first step in the chain, he is then taught to perform the first two behaviors in the chain, with the training completing the chain. This process is continued until the learner completes the entire chain independently.

fraudulent conduct Intentional, willful, and deceitful behavior; such behavior can cause harm to others.

free operant Any operant behavior that results in minimal displacement of the participant in time and space. A free operant can be emitted at nearly any time; it is discrete, it requires minimal time for completion, and it can produce a wide range of response rates. Examples in ABA include (a) the number of words read during a 1-min counting period, (b) the number of hand slaps per 6 seconds, and (c) the number of letter strokes written in 3 minutes. (Contrast with **discrete trial.**)

free-operant avoidance A contingency in which responses at any time during an interval prior to the scheduled onset of an aversive stimulus delays the presentation of the aversive stimulus. (Contrast with **discriminated avoidance.**)

frequency How often a behavior occurs. Some behavior analysts use *frequency* to mean *rate* (a ratio of responses per standard unit of time); others use *frequency* as a synonym for count.

full-session DRL A procedure for implementing DRL in which reinforcement is delivered at the end of the session if the total number of responses emitted during the session does not exceed a criterion limit. (See also **differential reinforcement of low rates [DRL].**)

function-altering effect (relevant to operant relations) A change in an organism's repertoire of MO, stimulus, and response relations, caused by reinforcement, punishment, an extinction procedure, or a recovery from punishment procedure. Respondent function-altering effects result from the pairing and unpairing of antecedent stimuli.

function-based definition Designates responses as members of the targeted response class solely in terms of their common effect on the environment.

functional analysis A term with two meanings in contemporary behavior analysis literature. In its original and most fundamental usage, *functional analysis* denotes demonstrations of functional relations between environmental variables and behavior. In the context of determining the purposes (functions) of problem behavior for an individual, *functional analysis* entails experimentally arranging antecedents and consequences representing those in the person's natural routines so that their separate effects on problem behavior can be observed and measured.

functional behavior assessment (FBA) A systematic method of assessment for obtaining information about the purposes (functions) a problem behavior serves for a person; results are used to guide the design of an intervention for decreasing the problem behavior and increasing appropriate behavior.

functional communication training (FCT) An antecedent intervention in which an appropriate communicative behavior is taught as a replacement behavior for problem behavior usually evoked by an establishing operation (EO); involves differential reinforcement of alternative behavior (DRA).

functional relation A verbal statement summarizing the results of an experiment (or group of related experiments) that describes the occurrence of the phenomena under study as a function of the operation of one or more specified and controlled variables in the experiment in which a specific change in one event (the dependent variable) can be produced by manipulating another event (the independent variable), and that the change in the dependent variable was unlikely the result of other factors (confounding variables); in behavior analysis expressed as $b = f(x1), (x2), \ldots$, where b is the behavior and $x1$, $x2$, etc., are environmental variables of which the behavior is a function.

functionally equivalent Serving the same function or purpose; different topographies of behavior are functionally equivalent if they produce the same consequences.

G

general case analysis A systematic process for identifying and selecting teaching examples that represent the full range of stimulus variations and response requirements in the generalization setting(s). (See also **multiple-exemplar training** and **teaching enough examples.**)

generalization A generic term for a variety of behavioral processes and behavior change outcomes. (See also **generalization gradient, generalized behavior change, response generalization, response maintenance, setting/situation generalization,** and **stimulus generalization.**)

generalization across subjects Changes in the behavior of people not directly treated by an intervention as a function of treatment contingencies applied to other people.

generalization probe Any measurement of a learner's performance of a target behavior in a setting and/or stimulus situation in which direct training has not been provided.

generalization setting Any place or stimulus situation that differs in some meaningful way from the instructional setting and in which performance of the target behavior is desired. (Contrast with **instructional setting.**)

generalized behavior change A behavior change that has not been taught directly. Generalized outcomes take one, or a combination of, three primary forms: response maintenance, stimulus/setting generalization, and response generalization. (Sometimes called *generalized outcome.*)

generalized conditioned reinforcer A conditioned reinforcer that as a result of having been paired with many other reinforcers does not depend on an establishing operation for any particular form of reinforcement for its effectiveness.

generalized imitation A term often used when a learner imitates a wide variety of unprompted, untrained, non-reinforced modeled behaviors in different settings and situations. Generalized imitation could be considered a higher order response class in which participants imitate novel models without training and that are not predictive of reinforcement.

generative learning A behavioral effect whereby previously acquired speaker and listener skills enable or accelerate the acquisition of other speaker and listener skills, without dependence on direct teaching or a history of reinforcement.

generic (tact) extension A tact evoked by a novel stimulus that shares all of the relevant or defining features associated with the original stimulus.

Good Behavior Game An interdependent group contingency in which a group is divided into two or more teams that compete against each other and/or a specified criterion. The team with fewest marks at the end of the game earns a reinforcer. Each team is also told that it can earn a reinforcer if it has fewer than a specified number of marks.

graph A visual format for displaying data; reveals relations among and between a series of measurements and relevant variables.

group contingency A contingency in which reinforcement for all members of a group is dependent on the behavior of (a) a person within the group, (b) a select group of members within the larger group, or (c) each member of the group meeting a performance criterion. (See also **dependent group contingency, independent group contingency,** and **interdependent group contingency.**)

H

habilitation Habilitation (adjustment) occurs when a person's repertoire has been changed such that short- and long-term reinforcers are maximized and short- and long-term punishers are minimized.

habit reversal A multiple-component treatment package for reducing unwanted habits such as fingernail biting and muscle tics; treatment typically includes self-awareness training involving response detection and procedures for identifying events that precede and trigger the response; competing response training; and motivation techniques including self-administered consequences, social support systems, and procedures for promoting the generalization and maintenance of treatment gains.

habituation A decrease in responsiveness to repeated presentations of a stimulus; most often used to describe a reduction of respondent behavior as a function of repeated presentation of the eliciting stimulus over a short span of time; some researchers suggest that the concept also applies to within-session changes in operant behavior.

hero procedure A term sometimes used for a dependent group contingency (i.e., a reward for the group is contingent upon the behavior of an individual group member).

hierarchical relation A nested stimulus relation in which a category, subsuming multiple stimuli, is itself a member of a higher-order category subsuming multiple stimuli.

high-probability (high-*p*) request sequence An antecedent intervention in which two to five easy tasks with a known history of learner compliance (the high-*p* requests) are presented in quick succession immediately before requesting the target task, the low-*p* request. (Also called *interspersed requests, pretask requests,* or *behavioral momentum.*)

high-probability instructional sequence (high-*p*). A procedure for improving compliance and diminishing escape-maintained problem behaviors. The practitioner presents two to five easy-to-follow (high-*p*) instructions for which the participant has a history of compliance. When the learner complies with several high-*p* instructions, the practitioner immediately presents a target instruction (i.e., low-*p*).

higher-order conditioning Development of a conditioned reflex by pairing of a neutral stimulus (NS) with a conditioned stimulus (CS). (Also called *secondary conditioning.*)

higher-order operant class Behavior defined in terms of general relations between antecedents and responses, rather than in terms of specific stimuli and responses; examples include generalized imitation, manding, instruction following, naming, and relational framing.

history of reinforcement An inclusive term referring in general to all of a person's learning experiences and more specifically to past conditioning with respect to particular response classes or aspects of a person's repertoire. (See also **ontogeny.**)

hypothetical construct A presumed but unobserved process or entity (e.g., Freud's id, ego, and superego).

I

imitation Behavior that (a) is occasioned by another person's model of the behavior (or symbolic representation thereof), (b) has formal similarity with the model, (c) follows the modeled behavior closely in time, and (d) the model is the primary controlling variable for the imitative behavior. An imitative behavior is a new behavior emitted following a novel antecedent event (i.e., the model). (See also **formal similarity;** contrast with **echoic.**)

imitation training A systematic, research-based set of procedures for teaching a nonimitative learner to imitate models of novel behaviors.

impure tact A verbal operant involving a response that is evoked by both an MO and a nonverbal stimulus; thus, the response is part mand and part tact. (See also **mand** and **tact.**)

independent group contingency A contingency in which reinforcement for each member of a group is dependent on that person's meeting a performance criterion that is in effect for all members of the group.

independent variable The variable that is systematically manipulated by the researcher in an experiment to see whether changes in the independent variable produce reliable changes in the dependent variable. In applied behavior analysis, it is usually an environmental event or condition antecedent or consequent to the dependent variable. (Sometimes called the *intervention* or *treatment variable*; compare to **dependent variable.**)

indirect functional assessment Structured interviews, checklists, rating scales, or questionnaires used to obtain information from people who are familiar with the person exhibiting the problem behavior (e.g., teachers, parents, caregivers, and/or the individual himself or herself); used to identify conditions or events in the natural environment that correlate with the problem behavior.

indirect measurement Occurs when the behavior that is measured is in some way different from the behavior of interest; considered less valid than direct measurement because inferences about the relation between the data obtained and the actual behavior of interest are required. (Contrast with **direct measurement.**)

indiscriminable contingency A contingency that makes it difficult for the learner to discriminate whether the next response will produce reinforcement. Practitioners use indiscriminable contingencies in the form of intermittent schedules of reinforcement and delayed rewards to promote generalized behavior change.

informed consent Formal permission—usually written, signed, and dated—given by a potential recipient of behavioral services or participant in a research study; requires that full disclosure of all pertinent information be provided to the individual prior to a granting of permission. To give consent, the person must (a) demonstrate the capacity to decide, (b) do so voluntarily, and (c) have adequate knowledge of all salient aspects of the treatment.

instructional setting The environment where instruction occurs; includes all aspects of the environment, planned and unplanned, that may influence the

learner's acquisition, maintenance, and generalization of the target behavior. (Contrast with **generalization setting.**)

interdependent group contingency A contingency in which reinforcement for all members of a group is dependent on each member of the group meeting a performance criterion that is in effect for all members of the group.

intermittent schedule of reinforcement (INT) A contingency of reinforcement in which some, but not all, occurrences of the behavior produce reinforcement.

internal validity The extent to which an experiment shows convincingly that changes in behavior are a function of the independent variable and not the result of uncontrolled or unknown variables. (Compare to **external validity.**)

interobserver agreement (IOA) The degree to which two or more independent observers report the same observed values after measuring the same events.

interresponse time (IRT) A measure of temporal locus, defined as the elapsed time between two successive responses.

interrupted chain procedure A procedural variation of the behavior chain interruption strategy, entails arranging the environment such that the learner is unable to continue the chain at a predetermined point until responding to a prompt (e.g., "What do you want?").

interval DRL A procedure for implementing DRL in which the total session is divided into equal intervals and reinforcement is provided at the end of each interval in which the number of responses during the interval is equal to or below a criterion limit. (See also **differential reinforcement of low rates [DRL].**)

interval-by-interval IOA An index of the agreement between observers for data obtained by interval recording or time sampling measurement; calculated for a given session or measurement period by comparing the two observers' recordings of the occurrence or nonoccurrence of the behavior in each observation interval and dividing the number of intervals of agreement by the total number of intervals and multiplying by 100. (Also called the *point-by-point* or *total interval IOA*; compare to **scored-interval IOA** and **unscored-interval IOA.**)

interview-informed synthesized contingency analysis A variation of functional analysis designed to increase efficiency. In the test condition, multiple contingencies are implemented simultaneously (e.g., attention and escape) when the problem behavior is demonstrated. In the control condition, those same contingencies are presented noncontingently and continuously.

intraverbal An elementary verbal operant involving a response that is evoked by a verbal discriminative stimulus that does *not* have point-to-point correspondence with that verbal stimulus. The intraverbal is the opposite of the echoic, in that the words emitted by one speaker do not match the words of another speaker. Intraverbal behavior constitutes the basis for social interaction, conversations, and much of academic and intellectual behavior. Questions are mands, and answers are intraverbal.

intraverbal control Some verbal stimuli only supplement other more critical antecedents, but nonetheless they play a causal role in evoking behavior. Palmer (2016) recommends "In cases in which the verbal antecedent is, by itself, insufficient to evoke the relevant response, we should speak of 'intraverbal control,' usually as one of a number of concurrent controlling variables" (p. 99).

irreversibility A situation that occurs when the level of responding observed in a previous phase cannot be reproduced even though the experimental conditions are the same as they were during the earlier phase.

J

joint control A phenomenon in which two separate, but interrelated forms of a person's own verbal behavior, combine to acquire stimulus control of a response that would not have occurred in the absence of either.

L

lag reinforcement schedule A schedule of reinforcement in which reinforcement is contingent on a response being different in some specified way (e.g., different topography) from the previous response (e.g., Lag 1) or a specified number of previous responses (e.g., Lag 2 or more).

latency See **response latency.**

latency-based functional analysis An analysis in which each session is terminated as soon as a problem behavior occurs. The index of problem behavior is the latency from onset of the establishing operation to the first occurrence of the problem behavior.

least-to-most response prompts A technique of transferring stimulus control in which the practitioner gives the participant an opportunity to perform the response with the least amount of assistance on each trial. The participant receives greater degrees of assistance with each successive trial without a correct response. The procedure for least-to-most prompting requires the participant to make a correct response within a set time limit (e.g., 3 seconds) from the presentation of the natural S^D. If the response does not occur within the specified time, the applied behavior analyst presents the natural S^D and a response prompt of least assistance, such as a verbal response prompt. If after the same specified time limit (e.g., another 3 seconds) the participant does not make a correct response, the analyst gives the natural S^D and another response prompt, such as a gesture. The participant receives partial or full physical guidance if the less intrusive prompt does not evoke a correct response.

level The value on the vertical axis around which a series of behavioral measures converge.

level system A component of some token economy systems in which participants advance up (or down) through a succession of levels contingent on their behavior at the current level. The performance criteria and sophistication or difficulty of the behaviors required at each level are higher than those of preceding levels; as participants advance to higher levels, they gain access to more desirable reinforcers, increased privileges, and greater independence.

limited hold A situation in which reinforcement is available only during a finite time following the elapse of an FI or VI interval; if the target response does not occur within the time limit, reinforcement is withheld and a new interval begins (e.g., on an FI 5-min schedule with a limited hold of 30 seconds, the first correct response following the elapse of 5 minutes is reinforced only if that response occurs within 30 seconds after the end of the 5-minute interval).

line graph Based on a Cartesian plane, a two-dimensional area formed by the intersection of two perpendicular lines. Any point within the plane represents a specific relation between the two dimensions described by the intersecting lines. It is the most common graphic format for displaying data in applied behavior analysis.

listener Someone who provides reinforcement for a speaker's verbal behavior. A listener may also serve as an audience evoking verbal behavior. The distinction between listener and speaker is often blurred by the fact that much of a listener's behavior may involve becoming a speaker at the covert level (e.g., thinking about what was said). A speaker may also serve as her own listener. (Contrast with **speaker.**)

listener discriminations When verbal S^D evokes a specific nonverbal behavior, due to a history of reinforcement.

local response rate The average rate of response during a smaller period of time within a larger period for which an overall response rate has been given. (See also **overall response rate.**)

M

magnitude The force or intensity with which a response is emitted; provides important quantitative parameters used in defining and verifying the occurrence of some response classes. Responses meeting those criteria are measured and reported by one or more fundamental or derivative measures such as frequency, duration, or latency. (Sometimes called *amplitude*.)

maintenance Two different meanings in applied behavior analysis: (a) the extent to which the learner continues to perform the target behavior after a portion or all of the intervention has been terminated (i.e., response maintenance), a dependent variable or characteristic of behavior; and (b) a condition in which treatment has been discontinued or partially withdrawn, an independent variable or experimental condition.

mand An elementary verbal operant involving a response of any form that is evoked by an MO and followed by specific reinforcement. Manding allows a speaker to get what she wants or refuse what she does not want.

massed practice A self-directed behavior change technique in which the person forces herself to perform an undesired behavior (e.g., a compulsive ritual) repeatedly, which sometimes decreases the future frequency of the behavior.

matching law The allocation of responses to choices available on concurrent schedules of reinforcement; rates of responding across choices are distributed in proportions that match the rates of reinforcement received from each choice alternative.

matching-to-sample procedure A discrete trial procedure for investigating conditional relations and stimulus equivalence. A matching-to-sample trial begins with the participant making a response that presents or reveals the sample stimulus; next, the sample stimulus may or may not be removed, and two or more comparison stimuli are presented. The participant then selects one of the comparison stimuli. Responses that select a comparison stimulus that matches the sample stimulus are reinforced; no reinforcement is provided for responses selecting the nonmatching comparison stimuli.

mean count-per-interval IOA The average percentage of agreement between the counts reported by two observers in a measurement period composed of a series of smaller counting times; a more conservative measure of IOA than total count IOA.

mean duration-per-occurrence IOA An IOA index for duration per occurrence data; also a more conservative and usually more meaningful assessment of IOA for total duration data calculated for a given session or measurement period by computing the average percentage of agreement of the durations reported by two observers for each occurrence of the target behavior.

measurement The process of applying quantitative labels to describe and differentiate objects and natural events. Measurement in applied behavior analysis involves three steps: (a) identifying the behavior to be measured, (b) defining the behavior in observable terms, and (c) selecting an appropriate observation and data-recording method.

measurement bias Nonrandom measurement error; a form of inaccurate measurement in which the data consistently overestimate or underestimate the true value of an event.

measurement by permanent product A method of measuring behavior after it has occurred by recording the effects that the behavior produced on the environment.

mentalism An approach to explaining behavior that assumes that a mental, or "inner," dimension exists that differs from a behavioral dimension and that phenomena in this dimension either directly cause or at least mediate some forms of behavior, if not all.

metaphorical (tact) extension A tact evoked by a novel stimulus that shares some, but not all, of the relevant features of the original stimulus.

methodological behaviorism A philosophical position that views behavioral events that cannot be publicly observed as outside the realm of science.

metonymical (tact) extension A tact evoked by a novel stimulus that shares none of the relevant features of the original stimulus configuration, but some irrelevant yet related feature has acquired stimulus control.

mixed schedule (mix) A compound schedule of reinforcement consisting of two or more basic schedules of reinforcement (elements) that occur in an alternating, usually random, sequence; no discriminative stimuli are correlated with the presence or absence of each element of the schedule, and reinforcement is delivered for meeting the response requirements of the element in effect at any time.

modeling A behavior change strategy in which learners acquire new skills by imitating demonstrations of the skills by live or symbolic models. The model shows, demonstrates, or conveys exactly the behavior the learner is expected to perform. Models can be live demonstrations or symbolic representations of the desired behavior.

momentary time sampling A measurement method in which the presence or absence of behaviors is recorded at precisely specified time intervals. (Contrast with **interval recording.**)

most-to-least response prompts A technique of transferring stimulus control in which the practitioner physically guides the participant through the entire performance sequence, and then gradually reduces the level of assistance in successive trials. Customarily, most-to-least prompting transitions from physical guidance to visual prompts to verbal instructions, and finally to the natural stimulus without prompts.

motivating operation (MO) An environmental variable that (a) alters (increases or decreases) the reinforcing or punishing effectiveness of some stimulus, object, or event; and (b) alters (increases or decreases) the current frequency of all behavior that has been reinforced or punished by that stimulus, object, or event. (See also **abative effect, abolishing operation [AO], behavior-altering effect, establishing operation [EO], evocative effect, value-altering effect.**)

motor imitation (relating to sign language) A type of duplic verbal behavior in which the form of a motor response is under the functional control of a visual verbal S^D that has formal similarity between a verbal stimulus and a verbal response product, and a history of generalized reinforcement.

multielement design An experimental design in which two or more conditions (one of which may be a no-treatment control condition) are presented in rapidly alternating succession (e.g., on alternating sessions or days) independent of the level of responding; differences in responding between or among conditions are attributed to the effects of the conditions. (Also called *alternating treatments design, concurrent schedule design*, and *multiple schedule design.*)

multiple baseline across behaviors design A multiple baseline design in which the treatment variable is applied to two or more different behaviors of the same subject in the same setting.

multiple baseline across settings design A multiple baseline design in which the treatment variable is applied to the same behavior of the same subject across two or more different settings, situations, or time periods.

multiple baseline across subjects design A multiple baseline design in which the treatment variable is applied to the same behavior of two or more subjects (or groups) in the same setting.

multiple baseline design An experimental design that begins with the concurrent measurement of two or more behaviors in a baseline condition, followed by the application of the treatment variable to one of the behaviors while baseline conditions remain in effect for the other behavior(s). After maximum change has been noted in the first behavior, the treatment variable is applied in sequential fashion to each of the other behaviors in the design. Experimental control is demonstrated if each behavior shows similar changes when, and only when, the treatment variable is introduced. (See also **multiple baseline across behaviors design**, **multiple baseline across settings design**, and **multiple baseline across subjects design**.)

multiple control (of verbal behavior) There are two types of multiple control. *Convergent multiple control* occurs when a single verbal response is a function of more than one variable (i.e., what is said has more than one antecedent source of control). *Divergent multiple control* occurs when a single antecedent variable affects the strength of more than one response.

multiple-exemplar training Instruction that provides the learner with practice with a variety of stimulus conditions, response variations, and response topographies to ensure the acquisition of desired stimulus control response forms; used to promote both setting/situation generalization and response generalization. (See also **teaching enough examples.**)

multiple probe design A variation of the multiple baseline design that features intermittent measures, or probes, during baseline. It is used to evaluate the effects of instruction on skill sequences in which it is unlikely that the subject can improve performance on later steps in the sequence before learning prior steps.

multiple schedule (mult) A compound schedule of reinforcement consisting of two or more basic schedules of reinforcement (elements) that occur in an alternating, usually random, sequence; a discriminative stimulus is correlated with the presence or absence of each element of the schedule, and reinforcement is delivered for meeting the response requirements of the element in effect at any time.

multiple treatment interference The effects of one treatment on a subject's behavior being confounding by the influence of another treatment administered in the same study.

multiple treatment reversal design Any experimental design that uses the experimental methods and logic of the reversal tactic to compare the effects of two or more experimental conditions to baseline and/or to one another (e.g., A-B-A-B-C-B-C, A-B-A-C-A-D-A-C-A-D, A-B-A-B-B+C-B-B+C).

mutual entailment A bidirectional stimulus relation in which one direction (e.g., if A, then B) is directly learned and the other (if B, then A) is derived.

N

naive observer An observer who is unaware of the study's purpose and/or the experimental conditions in effect during a given phase or observation period. Data obtained by a naive observer are less likely to be influenced by observers' expectations.

naturally existing contingency Any contingency of reinforcement (or punishment) that operates independent of the behavior analyst's or practitioner's efforts; includes socially mediated contingencies contrived by other people and already in effect in the relevant setting. (Contrast with **contrived contingency.**)

negative punishment A response behavior followed immediately by the removal of a stimulus (or a decrease in the intensity of the stimulus) that results in similar responses occurring less often. (Contrast with **positive punishment.**)

negative reinforcement A contingency in which the occurrence of a response is followed immediately by the termination, reduction, postponement, or avoidance of a stimulus, and which leads to an increase in the future occurrence of similar responses.

negative reinforcer A stimulus whose termination (or reduction in intensity) functions as reinforcement. (Contrast with **positive reinforcer.**)

negligence Failure to demonstrate professional integrity on the part of the provider that threatens the client's health and safety; typically exhibited as nonfeasance—not doing what ought to be done.

neutral stimulus (NS) A stimulus change that does not elicit respondent behavior. (Compare to **conditioned stimulus [CS]** and **unconditioned stimulus [US].**)

nodal stimulus (node) A stimulus set that must be held in common across a minimum of two conditional discriminations to provide a basis for all equivalence properties.

nonequivalence relations Derived stimulus relations in which stimuli are related on some basis other than "sameness."

noncontingent reinforcement (NCR) A procedure in which stimuli with known reinforcing properties are presented on fixed-time (FT) or variable-time (VT) schedules completely independent of behavior; often used as an antecedent intervention to reduce problem behavior. (See also **fixed-time schedule [FT], variable-time schedule [VT].**)

noncontingent reinforcement (NCR) reversal technique An experimental control technique that demonstrates the effects of reinforcement by using noncontingent reinforcement (NCR) as a control condition instead of a no-reinforcement (baseline) condition. During the NCR condition, the stimulus change used as reinforcement in the reinforcement condition is presented on a fixed or variable time schedule independent of the subject's behavior. A higher level of responding during the reinforcement condition than during the NCR condition demonstrates that the changes in behavior are the result of *contingent reinforcement*, not simply the presentation of or contact with the stimulus event. (Compare to **DRI/DRA reversal technique** and **DRO reversal technique.**)

nonconcurrent multiple baseline across participants design An experimental design that consists of a related series of A-B (baseline-intervention) sequences conducted across participants at different points in time; often used to assess treatment effects when concurrent measurement of participants' behavior is not possible.

nonexclusion time-out A procedure for implementing time-out in which, contingent on the occurrence of the target behavior, the person remains within the setting, but does not have access to reinforcement, for a specified period.

normalization As a philosophy and principle, the belief that people with disabilities should, to the maximum extent possible, be physically and socially integrated into the mainstream of society regardless of the degree or type of disability. As an approach to intervention, the use of progressively more typical settings and procedures "to establish and/or maintain personal behaviors which are as culturally normal as possible" (Wolfensberger, 1972, p. 28).

O

observed value A measure produced by an observation and measurement system. Observed values serve as the data that the researcher and others will interpret to form conclusions about an investigation. (Compare to **true value.**)

observer drift Any unintended change in the way an observer uses a measurement system over the course of an investigation that results in measurement error; often entails a shift in the observer's interpretation of the original definitions of the target behavior subsequent to being trained. (See also **measurement bias** and **observer reactivity.**)

observer reactivity Influence on the data reported by an observer that results from the observer's awareness that others are evaluating the data. (See also **measurement bias** and **observer drift.**)

ontogeny The history of the development of an individual organism during its lifetime. (See also **history of reinforcement;** compare to **phylogeny.**)

operant behavior Behavior that is selected, maintained, and brought under stimulus control as a function of its consequences; each person's repertoire of operant behavior is a product of his history of interactions with the environment (ontogeny).

operant conditioning The basic process by which operant learning occurs; consequences (stimulus changes immediately following responses) result in an increased (reinforcement) or decreased (punishment) frequency of the same type of behavior under similar motivational and environmental conditions in the future. (See also **motivating operation, punishment, reinforcement, response class,** and **stimulus control.**)

overall response rate The rate of response over a given time period. (See also **local response rate.**)

overcorrection A behavior change tactic based on positive punishment in which, contingent on the problem behavior, the learner is required to engage in effortful behavior directly or logically related to fixing the damage caused by the behavior. Forms of overcorrection are restitutional overcorrection and positive practice overcorrection. (See also **positive practice overcorrection** and **restitutional overcorrection.**)

overselective stimulus control A condition in which the range of discriminative stimuli, or stimulus features controlling behavior, is extremely limited; often interferes with learning. (Also called *stimulus overselectivity.*)

overshadowing Occurs when the most salient component of a compound stimulus arrangement controls responding and interferes with the acquisition of stimulus control by the more relevant stimulus.

P

parametric analysis An experiment designed to discover the differential effects of a range of values of an independent variable.

parsimony The practice of ruling out simple, logical explanations, experimentally or conceptually, before considering more complex or abstract explanations.

partial-interval recording A time sampling method for measuring behavior in which the observation period is divided into a series of brief time intervals (typically from 5 to 10 seconds). The observer records whether the target behavior occurred at any time during the interval. Partial-interval recording is not concerned with how many times the behavior occurred during the interval or how long the behavior was present, just that it occurred at some point during the interval; tends to overestimate the proportion of the observation period that the behavior actually occurred.

partition time-out A variation of time-out from positive reinforcement in which contingent on the occurrence of the target behavior, the participant remains in the time-in setting, but his view of the setting is restricted by a panel or cubicle, or a select space is arranged to serve as the time-out area (i.e., a carpet, a corner). Sometimes called *select space time-out.*

percentage A ratio (i.e., a proportion) formed by combining the same dimensional quantities, such as count (number ÷ number) or time (duration ÷ duration; latency ÷ latency); expressed as a number of parts per 100; typically expressed as a ratio of the number of responses of a certain type per total number of responses (or opportunities or intervals in which such a response could have occurred). A percentage presents a proportional quantity per 100.

perspective shifting Responding as if from the vantage point of another person, place, or time than the personal here and now.

philosophic doubt An attitude that the truthfulness and validity of all scientific theory and knowledge should be continually questioned.

phylogeny The history of the natural evolution of a species. (Compare to **ontogeny.**)

pivotal behavior A behavior that, when learned, produces corresponding modifications or covariation in other untrained behaviors. (Compare to **behavioral cusp.**)

placebo control A procedure that prevents a subject from detecting the presence or absence of the treatment variable. To the subject, the placebo condition appears the same as the treatment condition (e.g., a placebo pill contains an inert substance but looks, feels, and tastes exactly like a pill that contains the treatment drug). (See also **double-blind control.**)

planned activity check (PLACHECK) A variation of momentary time sampling in which the observer records whether each person in a group is engaged in the target behavior at specific points in time; provides a measure of "group behavior."

planned ignoring A procedure for implementing time-out in which social reinforcers—usually attention, physical contact, and verbal interaction—are withheld for a brief period contingent on the occurrence of the target behavior.

point-to-point correspondence A relation between the stimulus and response or response product that occurs when the beginning, middle, and end of the verbal stimulus matches the beginning, middle, and end of the verbal response. The verbal relations with point-to-point correspondence are echoic, copying text, imitation as it relates to sign language, textual, and transcription.

positive practice overcorrection A form of overcorrection in which, contingent on an occurrence of the target behavior, the learner is required to repeat a correct form of the behavior, or a behavior incompatible with the problem behavior, a specified number of times; entails an educative component. (See also **overcorrection** and **restitutional overcorrection.**)

positive punishment A response followed immediately by the presentation of a stimulus that decreases the future frequency of the behavior. (Contrast with **negative punishment.**)

positive reinforcement A response followed immediately by the presentation of a stimulus change that results in similar responses occurring more often. (Contrast with **negative reinforcement.**)

positive reinforcer A stimulus whose presentation or onset functions as reinforcement. (Contrast with **negative reinforcer.**)

postreinforcement pause The absence of responding for a period of time following reinforcement; an effect commonly produced by fixed interval (FI) and fixed ratio (FR) schedules of reinforcement.

practice effects Improvements in performance resulting from opportunities to perform a behavior repeatedly so that baseline measures can be obtained.

pragmatism A philosophical position asserting that the truth value of a statement is determined by how well it promotes effective action; pragmatism is a primary criterion by which behavior analysts judge the value of their findings.

precision teaching An instructional approach that involves (a) pinpointing the skills to be learned; (b) measuring the initial frequency or rate per minute at which the student can perform those skills; (c) setting an aim, or goal, for the child's improvement; (d) using direct, daily measurement to monitor progress made under an instructional program; (e) charting the results of those measurements on a Standard Celeration Chart; and (f) changing the program if progress is not adequate.

prediction A statement of the anticipated outcome of a presently unknown or future measurement; one of three components of the experimental reasoning, or baseline logic, used in single-subject research designs. (See also **replication** and **verification.**)

Premack principle A principle that states that making the opportunity to engage in a high-probability behavior contingent on the occurrence of a low-frequency behavior will function as reinforcement for the low-frequency behavior. Sometimes called "Grandma's Law." (See also **response-deprivation hypothesis.**)

principle of behavior A statement describing a functional relation between behavior and one or more of its controlling variables with generality across organisms, species, settings, behaviors, and time (e.g., extinction, positive reinforcement); an empirical generalization inferred from many experiments demonstrating the same functional relation.

private events Covert events typically accessible only to the person experiencing them. Skinner's radical behaviorism holds three major assumptions about private events: (a) private events are behavior; (b) behavior that takes place within the skin is distinguished from other ("public") behavior only by its inaccessibility; and (c) private behavior is influenced by (i.e., is a function of) the same kinds of variables as publicly accessible behavior.

procedural fidelity The extent to which procedures in all phases and conditions of an experiment, including baseline, are implemented correctly. (Contrast with **treatment integrity).**

programming common stimuli A tactic for promoting setting/situation generalization by making the instructional setting similar to the generalization setting; the two-step process involves (1) identifying salient stimuli that characterize the generalization setting and (2) incorporating those stimuli into the instructional setting.

progressive-ratio (PR) schedule of reinforcement A variation of the fixed ratio (FR) schedule of reinforcement that increases the ratio requirements incrementally within the session; PR schedule requirements are changed using (a) arithmetic progressions to add a constant number to each successive ratio, or (b) geometric progressions to add successively a constant proportion of the preceding ratio.

progressive schedule of reinforcement A schedule that systematically thins each successive reinforcement opportunity independent of the individual's behavior; progressive ratio (PR) and progressive interval (PI) schedules are thinned using arithmetic or geometric progressions.

progressive time delay A procedure for transferring stimulus control from contrived response prompts to naturally existing stimuli that starts with simultaneous presentation of the natural stimulus and the response prompt (i.e., 0-sec delay). The number of 0-sec trials depends on the task difficulty and the functioning level of the participant. Following the simultaneous presentations, the time delay is gradually and systematically extended.

punisher A stimulus change that decreases the future occurrence of behavior that immediately precedes it. (See also **aversive stimulus, conditioned punisher,** and **unconditioned punisher.**)

punishment A basic principle of behavior describing a response–consequence functional relation in which a response is followed immediately by a stimulus change that decreases future occurrences of that type of behavior.(See also **negative punishment** and **positive punishment.**)

R

radical behaviorism A form of behaviorism that attempts to understand all human behavior, including private events such as thoughts and feelings, in terms of controlling variables in the history of the person (ontogeny) and the species (phylogeny).

range-bound changing criterion design A variation of the changing criterion design in which each intervention sub-phase includes a lower and an upper criterion within which the participant is expected to perform.

rate A fundamental measure of how often behavior occurs expressed as count per standard unit of time (e.g., per minute, per hour, per day) and calculated by dividing the number of responses recorded by the number of standard units of time in which observations were conducted. Some behavior analysts use *rate* and *frequency* interchangeably; others use *frequency* to mean count. (See **frequency** and **count.**)

ratio scale A scale in which equal distances on a graph's axis correspond to equal ratios of change in the variable plotted on the axis.

ratio strain A behavioral effect associated with abrupt increases in ratio requirements when moving from denser to thinner reinforcement schedules; common effects include avoidance, aggression, and unpredictable pauses or cessation in responding.

reactivity Effects of an observation and measurement procedure on the behavior being measured. Reactivity is most likely when measurement procedures are obtrusive, especially if the person being observed is aware of the observer's presence and purpose.

recombinative generalization "Differential responding to novel combinations of stimulus components that have been included previously in other stimulus contexts" (Goldstein, 1983, p. 280).

recovery from punishment The occurrence of a previously punished type of response without its punishing consequence; analogous to the extinction of previously reinforced behavior and has the effect of undoing the effect of the punishment.

reflex A stimulus–response relation consisting of an antecedent stimulus and the respondent behavior it elicits (e.g., bright light–pupil contraction). Unconditioned and conditioned reflexes protect against harmful stimuli, help regulate the internal balance and economy of the organism, and promote reproduction. (See also **conditioned reflex, respondent behavior, respondent conditioning,** and **unconditioned reflex.**)

reflexive conditioned motivating operation (CMO-R) A stimulus that acquires MO effectiveness by preceding some form of worsening or improvement. It is exemplified by the warning stimulus in a typical escape–avoidance procedure, which establishes its own offset as reinforcement and evokes all behavior that has accomplished that offset.

reflexivity A type of stimulus-to-stimulus relation in which the learner, without any prior training or reinforcement for doing so, selects a comparison stimulus that is the same as the sample stimulus (e.g., A = A). Reflexivity would be demonstrated in the following matching-to-sample procedure: The sample stimulus is a picture of a tree, and the three comparison stimuli are a picture of a mouse, a picture of a cookie, and a duplicate of the tree picture used as the sample stimulus. The learner selects the picture of the tree without specific reinforcement in the past for making the tree-picture-to-tree-picture match. (Also called *generalized identity matching*; see also **stimulus equivalence;** compare to **transitivity** and **symmetry.**)

reinforcement A basic principle of behavior describing a response–consequence functional relation in which a response is followed immediately by a stimulus change that results in similar responses occurring more often. (See also **negative reinforcement** and **positive reinforcement.**)

reinforcer A stimulus change that increases the future frequency of behavior that immediately precedes it. (See also **positive reinforcer, conditioned reinforcer** and **unconditioned reinforcer.**)

reinforcer assessment Refers to a variety of direct, empirical methods for presenting one or more stimuli contingent on a target response and measuring their effectiveness as reinforcers.

reinforcer-abolishing effect (of a motivating operation) A decrease in the reinforcing effectiveness of a stimulus, object, or event caused by a motivating operation. For example, food ingestion abolishes (decreases) the reinforcing effectiveness of food.

reinforcer-establishing effect (of a motivating operation) An increase in the reinforcing effectiveness of a stimulus, object, or event caused by a motivating operation. For example, food deprivation establishes (increases) the reinforcing effectiveness of food.

relational frame Any specific type of arbitrarily applicable relational responding.

relational frame theory (RFT) A theory of derived stimulus relations proposing that stimulus relations are inherently verbal and that accumulated experience with relational exemplars creates generalized repertoires of relating.

relevance of behavior rule Holds that only behaviors likely to produce reinforcement in the person's natural environment should be targeted for change.

reliability (of measurement) Refers to the consistency of measurement, specifically, the extent to which repeated measurement of the same event yields the same values.

repeatability Refers to the fact that a behavior can occur repeatedly through time (i.e., behavior can be counted); one of the three dimensional quantities of behavior from which all behavioral measurements are derived. (See also **celeration, count, frequency, rate, temporal extent,** and **temporal locus.**)

repertoire All of the behaviors a person can do; or a set of behaviors relevant to a particular setting or task (e.g., gardening, mathematical problem solving).

replication (a) Repeating conditions within an experiment to determine the reliability of effects and increase internal validity. (See also **baseline logic, prediction,** and **verification.**) (b) Repeating whole experiments to determine the generality of findings of previous experiments to other subjects, settings, and/or behaviors. (See also **direct replication, external validity,** and **systematic replication.**)

research question A statement of what the researcher seeks to learn by conducting the experiment; may be presented in question form and is most often found in a published account as a statement of the experiment's purpose. All aspects of an experiment's design should follow from the research question.

resistance to extinction The relative frequency with which operant behavior is emitted during extinction.

respondent behavior The response component of a reflex; behavior that is elicited, or induced, by antecedent stimuli. (See also **reflex** and **respondent conditioning.**)

respondent conditioning A stimulus–stimulus pairing procedure in which a neutral stimulus (NS) is presented with an unconditioned stimulus (US) until the neutral stimulus becomes a conditioned stimulus that elicits the conditioned response. (Also called *classical* or *Pavlovian conditioning*; see also **conditioned reflex** and **higher order conditioning.**)

respondent extinction The repeated presentation of a conditioned stimulus (CS) in the absence of the unconditioned stimulus (US); the CS gradually loses its ability to elicit the conditioned response until the conditioned reflex no longer appears in the individual's repertoire.

response A single instance or occurrence of a specific class or type of behavior. Technical definition: an "action of an organism's effector. An effector is an organ at the end of an efferent nerve fiber that is specialized for altering its environment mechanically, chemically, or in terms of other energy changes" (Michael, 2004, p. 8). (See also **response class.**)

response blocking A procedure in which the therapist physically intervenes as soon as the learner begins to emit a problem behavior, to prevent completion of the targeted behavior.

response class A group of responses of varying topography, all of which produce the same effect on the environment.

response cost The response-contingent loss of a specific number of positive reinforcers (e.g., a fine) that decreases the frequency of similar responses in the future; a form of negative punishment.

response differentiation A behavior change produced by differential reinforcement: Reinforced members of the current response class occur with greater frequency, and unreinforced members occur less frequently (undergo extinction); the overall result is the emergence of a new response class.

response generalization The extent to which a learner emits untrained responses that are functionally equivalent to the trained target behavior. (Compare to **response maintenance** and **setting/situation generalization.**)

response interruption and redirection (RIRD) A procedural variation of response blocking that involves interrupting stereotypic behavior at its onset and redirecting the individual to complete high-probability behaviors instead.

response latency A measure of temporal locus; the elapsed time from the onset of a stimulus (e.g., task direction, cue) to the initiation of a response.

response maintenance The extent to which a learner continues to perform the target behavior after a portion or all of the intervention responsible for the behavior's initial appearance in the learner's repertoire has been terminated. (Often called *maintenance, durability, behavioral persistence*, and (incorrectly) *resistance to extinction*; compare to **response generalization** and **setting/situation generalization.**)

response prompts Prompts that operate directly on the response to cue a correct response. The three major forms of response prompts are verbal instructions, modeling, and physical guidance.

response-deprivation hypothesis A model for predicting whether contingent access to one behavior will function as reinforcement for engaging in another behavior based on whether access to the contingent behavior represents a restriction of the activity compared to the baseline level of engagement. (See also **Premack principle.**)

restitutional overcorrection A form of overcorrection in which, contingent on the problem behavior, the learner is required to repair the damage or return the environment to its original state and then to engage in additional behavior

to bring the environment to a condition vastly better than it was in prior to the misbehavior. (See also **overcorrection** and **positive practice overcorrection.**)

restraint Three common types of restraint—personal restraint, protective equipment restraint, and self-restraint—appear in the behavior analysis literature. With *personal restraint*, therapists physically restrict the problem behavior by holding the body parts (e.g., response blocking) that produce the problem behavior. With *protective equipment restraints*, therapists often use protective equipment restraints (e.g., padded helmets, arm splints, boxing gloves, padded hand mitts, safety belts, safety straps) to protect individuals from causing serious injury. Persons with self-injurious behavior (SIB) often apply *self-restraint*, such as sitting on hands, placing arms between folded legs, and wrapping arms in tight clothing.

resurgence Refers to the reoccurrence of a previously reinforced behavior when reinforcement for an alternative behavior is terminated or decreased and to the three-phase procedure that produces the effect: (1) A target behavior is reinforced, (2) the target behavior is placed on extinction and reinforcement provided for an alternative behavior, and (3) both responses are placed on extinction.

reversal design Any experimental design in which the researcher attempts to verify the effect of the independent variable by "reversing" responding to a level obtained in a previous condition; encompasses experimental designs in which the independent variable is withdrawn (A-B-A-B) or reversed in its focus (e.g., DRI/DRA). (See also **A-B-A design, A-B-A-B design, B-A-B design, DRI/DRA reversal technique, DRO reversal technique,** and **noncontingent reinforcement [NCR] reversal technique.**)

risk–benefit analysis A comparison prior to conducting a behavioral assessment or intervention of the potential harm to the client (or others) versus the benefits likely to result from the procedures.

rule-governed behavior Behavior controlled by a rule (i.e., a verbal statement of an antecedent-behavior-consequence contingency); enables human behavior (e.g., fastening a seat belt) to come under the indirect control of temporally remote or improbable, but potentially significant consequences (e.g., avoiding injury in an auto accident). Often used in contrast to *contingency-shaped behavior*, a term used to indicate behavior selected and maintained by controlled, temporally close consequences.

S

satiation A decrease in the frequency of operant behavior presumed to be the result of continued contact with or consumption of a reinforcer that has followed the behavior; also refers to a procedure for reducing the effectiveness of a reinforcer (e.g., presenting a person with copious amounts of a reinforcing stimulus prior to a session). (See also **motivating operation;** contrast with **deprivation.**)

scatterplot A two-dimensional graph that shows the relative distribution of individual measures in a data set with respect to the variables depicted by the *x*- and *y*-axes. Data points on a scatterplot are not connected.

scatterplot recording A procedure for recording the extent to which a target behavior occurs more often at particular times than others; involves dividing the day into blocks of time, and using different symbols on an observation form to indicate the level of the target behavior (e.g., a lot, some, or not at all).

schedule of reinforcement A rule specifying the environmental arrangements and response requirements for reinforcement; a description of a contingency of reinforcement.

schedule thinning Changing a contingency of reinforcement by gradually increasing the response ratio or the extent of the time interval; it results in a lower rate of reinforcement per responses, time, or both.

science A systematic approach to the understanding of natural phenomena (as evidenced by description, prediction, and control) that relies on determinism as its fundamental assumption, empiricism as its primary rule, experimentation as its basic strategy, replication as a requirement for believability, parsimony as a value, and philosophic doubt as its guiding conscience.

scored-interval IOA An interobserver agreement index based only on the intervals in which either observer recorded the occurrence of the behavior; calculated by dividing the number of intervals in which the two observers agreed that the behavior occurred by the number of intervals in which either or both observers recorded the occurrence of the behavior and multiplying by 100. Scored-interval IOA is recommended as a measure of agreement for behaviors that occur at low rates because it ignores the intervals in which agreement by chance is highly likely. (Compare to **interval-by-interval IOA** and **unscored-interval IOA.**)

selection by consequences The fundamental principle underlying operant conditioning; the basic tenet is that all forms of (operant) behavior, from simple to complex, are selected, shaped, and maintained by their consequences during an individual's lifetime; Skinner's concept of selection by consequences is parallel to Darwin's concept of natural selection of genetic structures in the evolution of species.

selection-based verbal behavior A category of verbal behavior in which the speaker points to or selects a particular stimulus; what is conveyed to the listener is the information on the stimulus selected. (Contrast with **topography-based verbal behavior**)

selectionism A theory that all forms of life naturally and continually evolve as a result of the interaction between function and the survival value of that function. Operant selection by consequences is the conceptual and empirical foundation of behavior analysis.

self-contract Contingency contract that a person makes with himself or herself, incorporating a self-selected task and reward as well as personal monitoring of task completions and self-delivery of the reward.

self-control [impulse analysis] A person's ability to "delay gratification" by emitting a response that will produce a larger (or higher quality) delayed reward over a response that produces a smaller but immediate reward. (Sometimes called *impulse control.*)

self-control [Skinner's analysis] Skinner (1953) conceptualized self-control as a two-response phenomenon: The *controlling response* affects variables in such a way as to change the probability of the *controlled response*. (See also **self-management.**)

self-evaluation A procedure in which a person compares her performance of a target behavior with a predetermined goal or standard; often a component of self-management. (Sometimes called *self-assessment.*)

self-instruction Self-generated verbal responses, covert or overt, that function as rules or response prompts for a desired behavior; as a self-management tactic, self-instruction can guide a person through a behavior chain or sequence of tasks.

self-management The personal application of behavior change tactics that produces a desired change in behavior.

self-monitoring A procedure whereby a person systematically observes his behavior and records the occurrence or nonoccurrence of a target behavior. (Also called *self-recording* or *self-observation.*)

semilogarithmic chart A two-dimensional graph with a logarithmic scaled *y* axis so that equal distances on the vertical axis represent changes in behavior that are of equal proportion. (See also **Standard Celeration Chart.**)

sensory extinction A procedure by which behaviors maintained by automatic reinforcement are placed on extinction by masking or removing the sensory consequence.

sequence effects The effects on a subject's behavior in a given condition that are the result of the subject's experience with a prior condition.

setting/situation generalization The extent to which a learner emits the target behavior in a setting or stimulus situation that is different from the instructional setting.

shaping Using differential reinforcement to produce a series of gradually changing response classes; each response class is a successive approximation toward a terminal behavior. Members of an existing response class are selected for differential reinforcement because they more closely resemble the terminal behavior. (See also **differential reinforcement, response class, response differentiation,** and **successive approximations.**)

simple discrimination Responding is under stimulus control of a single antecedent stimulus condition; described by the three-term contingency: $S^D \rightarrow R \rightarrow S^{R+}$.

simple verbal discrimination A single-component word or phrase evokes a nonmatching intraverbal response (e.g., upon hearing "Ready, set . . . ," a child says "go,")

simple-to-complex testing protocol An arrangement in which types of stimulus equivalence probes are introduced sequentially, beginning with symmetry, followed by transitivity (if relevant), and then combined tests for equivalence.

single-case designs A wide variety of research designs that use a form of experimental reasoning called *baseline logic* to demonstrate the effects of the independent variable on the behavior of individual subjects. (Also called *single-subject, within-subject*, and *intra-subject* designs; see also **baseline logic, changing criterion design, multielement design, multiple baseline design, reversal design,** and **steady state strategy.**)

social validity Refers to the extent to which target behaviors are appropriate, intervention procedures are acceptable, and important and significant changes in target and collateral behaviors are produced.

socially mediated contingency (reinforcement) A contingency in which an antecedent stimulus and/or the consequence for the behavior is presented by another person.

solistic (tact) extension A verbal response evoked by a stimulus property that is only indirectly related to the proper tact relation (e.g., Yogi Berra's classic malapropism: "Baseball is ninety percent mental; the other half is physical."

spaced-responding DRL A procedure for implementing DRL in which reinforcement follows each occurrence of the target behavior that is separated from the previous response by a minimum interresponse time (IRT). (See also **differential reinforcement of low rates [DRL].**)

spatial relation Responding jointly to two stimuli on the basis of their juxtaposition in space.

speaker Someone who engages in verbal behavior by emitting mands, tacts, intraverbals, autoclitics, etc. A speaker is also someone who uses sign language, gestures, signals, written words, codes, pictures, or any form of verbal behavior. (Contrast with **listener.**)

split-middle line of progress A line drawn through a series of graphed data points that shows the overall trend in the data; drawn through the intersections of the vertical and horizontal middles of each half of the charted data and then adjusted up or down so that half of all the data points fall on or above and half fall on or below the line.

spontaneous recovery A behavioral effect associated with extinction in which the behavior suddenly begins to occur after its frequency has decreased to its prereinforcement level or stopped entirely.

stable baseline Data that show no evidence of an upward or downward trend; all of the measures fall within a relatively small range of values. (See also **steady state responding.**)

Standard Celeration Chart A multiply–divide chart with six base-10 (or × 10, ÷ 10) cycles on the vertical axis that can accommodate response rates as low as 1 per 24 hours (0.000695 per minute) to as high as 1000 per minute. It enables the standardized charting of celeration, a factor by which rate of behavior multiplies or divides per unit of time. (See also **semilogarithmic chart.**)

steady state responding A pattern of responding that exhibits relatively little variation in its measured dimensional quantities over a period of time.

steady state strategy Repeatedly exposing a subject to a given condition while trying to eliminate or control extraneous influences on the behavior and obtaining a stable pattern of responding before introducing the next condition.

stimulus "An energy change that affects an organism through its receptor cells" (Michael, 2004, p. 7).

stimulus blocking Occurs when a competing stimulus blocks the evocative function of a stimulus that has acquired stimulus control over the behavior. (Sometimes called *masking*.) Stimulus blocking can be mitigated by rearranging the physical environment, making instructional stimuli appropriately intense, and consistently reinforcing behavior in the presence of the instructionally relevant discriminative stimuli.

stimulus class A group of stimuli that share specified common elements along formal (e.g., size, color), temporal (e.g., antecedent or consequent), and/or functional (e.g., discriminative stimulus) dimensions.

stimulus control A situation in which the frequency, latency, duration, or amplitude of a behavior is altered by the presence or absence of an antecedent stimulus. (See also **discrimination** and **discriminative stimulus.**)

stimulus delta (S^{Δ}) A stimulus in the presence of which a given behavior has not produced reinforcement, or has produced reinforcement of lesser quality, in the past. (Contrast with **discriminative stimulus [S^D].**)

stimulus discrimination When one stimulus (the S^D) signals the availability of reinforcement and the absence of that stimulus (the S^{Δ}) signals a zero or reduced chance of reinforcement, responses will occur more often in the presence of the S^D than in its absence (the S^{Δ}).

stimulus discrimination training The conventional procedure requires one behavior and two antecedent stimulus conditions. Responses are reinforced in the presence of one stimulus condition, the S^D, but not in the presence of the other stimulus, the S^{Δ}.

stimulus equivalence The emergence of accurate responding to untrained and nonreinforced stimulus–stimulus relations following the reinforcement of responses to some stimulus–stimulus relations. Requires successful performances on three types of probe trials—reflexivity, symmetry, and transitivity—in the absence of reinforcement. (Also called *equivalence-class formation*.)

stimulus fading A method of transferring stimulus control that involves highlighting a physical dimension of a stimulus (e.g., size, color, position) to increase the likelihood of a correct response and then gradually diminishing the exaggerated dimension until the learner is responding correctly to the naturally occurring stimulus.

stimulus generalization When an antecedent stimulus has a history of evoking a response that has been reinforced in its presence, the same type of behavior tends to be evoked by stimuli that share similar physical properties with the controlling antecedent stimulus.

stimulus generalization gradient A graphic depiction of the extent to which behavior that has been reinforced in the presence of a specific stimulus condition is emitted in the presence of other stimuli. The gradient shows relative degrees of stimulus generalization and stimulus control (or discrimination). A flat slope across test stimuli shows a high degree of stimulus generalization and relatively little discrimination between the trained stimulus and other stimuli; a slope that drops sharply from its highest point corresponding to the trained stimulus indicates a high degree of stimulus control (discrimination) and relatively little stimulus generalization.

stimulus preference assessment A variety of procedures used to determine the stimuli that a person prefers, the relative preference values (high versus low) of those stimuli, the conditions under which those preference values remain in effect, and their presumed value as reinforcers.

stimulus prompts Prompts that operate directly on the antecedent task stimuli to cue a correct response in conjunction with the critical S^D (e.g., changing the size, color, or position of a stimulus within an array to make its selection more likely).

stimulus–control topographies Refers to two different forms of stimulus control that can result from a match-to-sample procedure involving one sample stimulus and two comparison stimuli; when presented with A1 as the sample, a participant can either *select* B1 (known as *Type S* or *select responding*) or *reject* B2 (known as *Type R* or *reject responding*).

stimulus–stimulus pairing A procedure in which two stimuli are presented at the same time, usually repeatedly for a number of trials, which often results in one stimulus acquiring the function of the other stimulus.

successive approximations The sequence of new response classes that emerge during the shaping process as the result of differential reinforcement; each successive response class is closer in form to the terminal behavior than the response class it replaces.

surrogate conditioned motivating operation (CMO-S) A stimulus that acquires its MO effectiveness by being paired with another MO and has the same value-altering and behavior-altering effects as the MO with which it was paired.

symmetry A type of stimulus-to-stimulus relationship in which the learner, without prior training or reinforcement for doing so, demonstrates the reversibility of matched sample and comparison stimuli (e.g., if A = B, then B = A). Symmetry would be demonstrated in the following matching-to-sample procedure: The learner is taught, when presented with the spoken word *car* (sample stimulus A), to select a comparison picture of a car (comparison B).

When presented with the picture of a car (sample stimulus B), without additional training or reinforcement, the learner selects the comparison spoken word *car* (comparison A). (See also **stimulus equivalence**; compare to **reflexivity** and **transitivity.**)

systematic desensitization A behavior therapy treatment for anxieties, fears, and phobias that involves substituting one response, generally muscle relaxation, for the unwanted behavior—the fear and anxiety. The client practices relaxing while imagining anxiety-producing situations in a sequence from the least fearful to the most fearful.

systematic replication An experiment in which the researcher purposefully varies one or more aspects of an earlier experiment. A systematic replication that reproduces the results of previous research not only demonstrates the reliability of the earlier findings but also adds to the external validity of the earlier findings by showing that the same effect can be obtained under different conditions.

T

tact An elementary verbal operant involving a response that is evoked by a nonverbal discriminative stimulus and followed by generalized conditioned reinforcement. Tacting allows a speaker to identify or describe the features of the physical environment. The elements that make up one's physical environment are vast; thus, much of language instruction and educational programs focus on teaching tacts.

tact extension Once a tact has been established, the tact response can occur under novel stimulus conditions through the process of stimulus generalization. Skinner (1957) identifies four different levels of generalization based on the degree to which a novel stimulus shares the relevant or defining features of the original stimulus. The four types of tact extension are generic, metaphorical, metonymical, and solecistic.

taking dictation An elementary verbal operant involving a spoken verbal stimulus that evokes a written, typed, or fingerspelled response that does not have formal similarity between the stimulus and the response, but does have point-to-point correspondence and a history of generalized reinforcement.

tandem schedule (tand) A schedule of reinforcement identical to the chained schedule except, like the mix schedule, the tandem schedule does not use discriminative stimuli with the elements in the chain. (See also **chained schedule** and **mixed schedule.**)

target behavior The response class selected for intervention; can be defined either functionally or topographically.

task analysis The process of breaking a complex skill or series of behaviors into smaller, teachable units; also refers to the results of this process.

teach loosely Randomly varying functionally irrelevant stimuli within and across teaching sessions; promotes setting/situation generalization by reducing the likelihood that (a) a single or small group of noncritical stimuli will acquire exclusive control over the target behavior and (2) the learner's performance of the target behavior will be impeded or "thrown off" should he encounter any of the "loose" stimuli in the generalization setting.

teach enough examples A strategy for promoting generalized behavior change that consists of teaching the learner to respond to a subset of all the relevant stimulus and response examples and then assessing the learner's performance on untrained examples. (See also **multiple exemplar training.**)

temporal extent Refers to the fact that every instance of behavior occurs during some amount of time; one of the three dimensional quantities of behavior from which all behavioral measurements are derived. (See also **repeatability** and **temporal locus.**)

temporal locus Refers to the fact that every instance of behavior occurs at a certain point in time with respect to other events (i.e., when in time behavior occurs can be measured); often measured in terms of *response latency* and *interresponse time (IRT)*; one of the three dimensional quantities of behavior from which all behavioral measurements are derived. (See also **repeatability** and **temporal extent.**)

temporal relation Responding jointly to two stimuli on the basis of their juxtaposition in time.

terminal behavior The end product of shaping.

terminate specific reinforcer contact A variation of nonexclusion time-out whereby each occurrence of the target behavior immediately stops an activity or sensory reinforcer.

textual An elementary verbal operant involving a response that is evoked by a written verbal discriminative stimulus that does not have formal similarity between the stimulus and the response, but does have point-to-point correspondence and a history of generalized reinforcement.

three-term contingency The basic unit of analysis in the analysis of operant behavior; encompasses the temporal and possibly dependent relations among an antecedent stimulus, behavior, and consequence.

time delay A procedure for transferring stimulus control from contrived response prompts to naturally existing stimuli that begins with the simultaneous presentation of the natural stimulus and response prompt. After several correct responses, a delay is introduced between the stimulus and the response prompt until the student emits the unprompted correct response. Time delay is considered an "errorless learning" technique as students make few or no errors transitioning from the contrived prompt to the instructional stimulus. (See also **constant time delay** and **progressive time delay.**)

time sampling A measurement of the presence or absence of behavior within specific time intervals. It is most useful with continuous and high-rate behaviors. (See also **momentary time sampling, partial-interval recording,** and **whole-interval recording.**)

time-out from positive reinforcement The immediate response-contingent withdrawal of the opportunity to earn positive reinforcement or the immediate loss of access to positive reinforcers for a specified time; a form of negative punishment. (Also called *time-out.*)

token An object or symbol that is awarded contingent on appropriate target behavior(s) that can be traded for a wide variety of backup reinforcers; tokens function as generalized conditioned reinforcers.

token economy A behavior change system consisting of a list of target behaviors, with tokens (points or small objects) participants earn for emitting the target behaviors, and a menu of backup reinforcers (i.e., preferred items, activities, or privileges) for which participants exchange earned tokens. (Also called *token reinforcement system.*)

topography The physical form or shape of a behavior.

topography-based definition Defines instances of the targeted response class by the shape or form of the behavior.

topography-based verbal behavior A category of verbal behavior in which the listener is affected by a specific response topography emitted by the speaker; includes (e.g., speech, sign language, writing, fingerspelling). (Contrast with **selection-based verbal behavior**)

total count IOA The simplest indicator of IOA for event recording data; based on comparing the total count recorded by each observer per measurement period; calculated by dividing the smaller of the two observers' counts by the larger count and multiplying by 100.

total duration IOA A relevant index of IOA for total duration measurement; computed by dividing the shorter of the two durations reported by the observers by the longer duration and multiplying by 100.

total-task chaining A variation of forward chaining in which the learner receives training on each behavior in the chain during each session.

training structure Refers to dimensions of procedural arrangements when teaching multiple conditional discriminations. Commonly used training structures include *one-to-many training* or *sample-as-node training*, *many-to-one* or *comparison-as-node structure*, and *linear series training*.

transcription An elementary verbal operant involving a spoken verbal stimulus that evokes a written, typed, or finger-spelled response. Like the textual, there is point-to-point correspondence between the stimulus and the response product, but no formal similarity.

transfer of function Occurs when teaching a new function for one member of an established equivalence class results in the same function holding for all members of the class.

transformation of function Occurs when the behavioral function of one stimulus in a stimulus class changes as a predictable function of the behavior function of other stimuli in the class.

transitive conditioned motivating operation (CMO-T) An environmental variable that, as a result of a learning history, establishes (or abolishes) the reinforcing effectiveness of another stimulus and evokes (or abates) the behavior that has been reinforced by that other stimulus.

transitivity Describes derived stimulus–stimulus relations (e.g., A = C) that emerge as a product of training two other stimulus–stimulus relations (A = B and B = C). Transitivity would be demonstrated if, after training, the following occurs: (1) If A (e.g., spoken word *bicycle*) = B (e.g., the picture of a bicycle) and (2) B (the picture of a bicycle) = C (e.g., the written word *bicycle*), then (3) A (the spoken name, *bicycle*) = C (the written word *bicycle*). (See also **stimulus equivalence;** compare to **reflexivity** and **symmetry**).

treatment drift An undesirable situation in which the independent variable of an experiment is applied differently during later stages than it was at the outset of the study.

treatment integrity The extent to which the independent variable is applied exactly as planned and described and no other unplanned variables are administered inadvertently along with the planned treatment. (Contrast with **procedural fidelity.**)

treatment package A behavioral intervention consisting of multiple components (e.g., contingent praise, tokens, and extinction).

trend The overall direction taken by a data path. It is described in terms of direction (increasing, decreasing, or zero trend), degree (gradual or steep), and the extent of variability of data points around the trend. Trend is used in predicting future measures of the behavior under unchanging conditions.

trial-based functional analysis An analysis in which a series of trials is interspersed among classroom activities. Each trial consists of two 1-minute components: (a) the establishing operation and contingency for problem behavior (test condition), and (b) continuous access to the reinforcer (control condition).

trial-by-trial IOA An IOA index for discrete trial data based on comparing the observers' counts (0 or 1) on a trial-by-trial, or item-by-item, basis; yields a more conservative and meaningful index of IOA for discrete trial data than does total count IOA.

true value A measure accepted as a quantitative description of the true state of some dimensional quantity of an event as it exists in nature. Obtaining true values requires "special or extraordinary precautions to ensure that all possible sources of error have been avoided or removed" (Johnston & Pennypacker, 1993a, p. 136). (Compare to **observed value.**)

trials-to-criterion A special form of event recording; a measure of the number of responses or practice opportunities needed for a person to achieve a pre-established level of accuracy or proficiency.

Type I error An error that occurs when a researcher concludes that the independent variable had an effect on the dependent variable, when no such relation exists; a *false positive*. (Contrast with **Type II error.**)

Type II error An error that occurs when a researcher concludes that the independent variable had no effect on the dependent variable, when in truth it did; a *false negative*. (Contrast with **Type I error.**)

U

unchaining Occurs when, in the case of a two-step chain, the second behavior in the chain (R_2) produces reinforcement in the presence of the prior S^D (S_2), but (R_2) also produces reinforcement when that S^D is not present. Unchaining may weaken a chain.

unconditioned motivating operation (UMO) A motivating operation whose value-altering effect does not depend on a learning history. For example, food deprivation increases the reinforcing effectiveness of food without the necessity of any learning history.

unconditioned negative reinforcer A stimulus that functions as a negative reinforcer as a result of the evolutionary development of the species (phylogeny); no prior learning is involved (e.g., shock, loud noise, intense light, extreme temperatures, strong pressure against the body). (See also **negative reinforcer;** compare to **conditioned negative reinforcer.**)

unconditioned punisher A stimulus change that decreases the frequency of any behavior that immediately precedes it irrespective of the organism's learning history with the stimulus. Unconditioned punishers are products of the evolutionary development of the species (phylogeny), meaning that all members of a species are more or less susceptible to punishment by the presentation of unconditioned punishers. (Also called *primary* or *unlearned punishers*; compare to **conditioned punisher.**)

unconditioned reflex An unlearned stimulus–response functional relation consisting of an antecedent stimulus (e.g., food in mouth) that elicits the response (e.g., salivation); a product of the phylogenic evolution of a given species; all biologically intact members of a species are born with similar repertoires of unconditioned reflexes. (See also **conditioned reflex.**)

unconditioned reinforcer A stimulus change that increases the frequency of any behavior that immediately precedes it irrespective of the organism's learning history with the stimulus. Unconditioned reinforcers are the product of the evolutionary development of the species (phylogeny). (Also called *primary* or *unlearned reinforcer*; compare to **conditioned reinforcer.**)

unconditioned stimulus (US) The stimulus component of an unconditioned reflex; a stimulus change that elicits respondent behavior without any prior learning.

unpairing Two kinds: (a) The occurrence alone of a stimulus that acquired its function by being paired with an already effective stimulus, or (b) the occurrence of the stimulus in the absence as well as in the presence of the effective stimulus. Both kinds of unpairing undo the result of the pairing: the occurrence alone of the stimulus that became a conditioned reinforcer; and the occurrence of the unconditioned reinforcer in the absence as well as in the presence of the conditioned reinforcer.

unscored-interval IOA An interobserver agreement index based only on the intervals in which either observer recorded the nonoccurrence of the behavior; calculated by dividing the number of intervals in which the two observers agreed that the behavior did not occur by the number of intervals in which either or both observers recorded the nonoccurrence of the behavior and multiplying by 100. Unscored-interval IOA is recommended as a measure of agreement for behaviors that occur at high rates because it ignores the intervals in which agreement by chance is highly likely. (Compare to **interval-by-interval IOA** and **scored-interval IOA.**)

V

validity (of measurement) The extent to which data obtained from measurement are directly relevant to the target behavior of interest and to the reason(s) for measuring it.

value-altering effect (of a motivating operation) Either (a) an increase in the reinforcing effectiveness of some stimulus, object, or event, in which case the MO is an establishing operation (EO); or (b) a decrease in reinforcing effectiveness, in which case the MO is an abolishing operation (AO). For example, the reinforcing effectiveness of food is altered as a result of food deprivation and food ingestion.

variability The frequency and extent to which multiple measures of behavior yield different outcomes.

variable baseline Data points that do not consistently fall within a narrow range of values and do not suggest any clear trend.

variable interval (VI) A schedule of reinforcement that provides reinforcement for the first correct response following the elapse of variable durations of time occurring in a random or unpredictable order. The mean duration of the intervals is used to describe the schedule (e.g., on a VI 10-min schedule, reinforcement is delivered for the first response following an average of 10 minutes since the last reinforced response, but the time that elapses following the last reinforced response might range from 30 seconds or less to 25 minutes or more).

variable ratio (VR) A schedule of reinforcement requiring a varying number of responses for reinforcement. The number of responses required varies around a random number; the mean number of responses required for reinforcement is used to describe the schedule (e.g., on a VR 10 schedule an

average of 10 responses must be emitted for reinforcement, but the number of responses required following the last reinforced response might range from 1 to 30 or more).

variable-interval DRO (VI-DRO) A DRO procedure in which reinforcement is available at the end of intervals of variable duration and delivered contingent on the absence of the problem behavior during the interval. (See also **differential reinforcement of other behavior [DRO].**)

variable-momentary DRO (VM-DRO) A DRO procedure in which reinforcement is available at specific moments of time, which are separated by variable amounts of time in random sequence, and delivered if the problem is not occurring at those times. (See **also differential reinforcement of other behavior [DRO].**)

variable-time schedule (VT) A schedule for the delivery of noncontingent stimuli in which the interval of time from one delivery to the next randomly varies around a given time. For example, on a VT 1-min schedule, the delivery-to-delivery interval might range from 5 seconds to 2 minutes, but the average interval would be 1 minute.

verbal behavior Behavior whose reinforcement is mediated by a listener; includes both vocal-verbal behavior (e.g., saying "*Water, please*" to get water) and nonvocal-verbal behavior (pointing to a glass of water to get water). Encompasses the subject matter usually treated as language and topics such as thinking, grammar, composition, and understanding.

verbal conditional discrimination (VC^D) A type of convergent multiple control involving a verbal stimulus that alters the evocative effects of another verbal stimulus in the same antecedent configuration. The conditional discrimination is between the words in the antecedent event.

verbal episode An interaction between a speaker and a listener. A speaker emits any type of verbal response (e.g., echoic, mand, tact, intraverbal), in any form (speech, sign language, icon selection, eye contact), and a listener (1) serves as an audience for a speaker, (2) provides reinforcement for a speaker, and (3) responds in specific ways to the speaker's behavior. The roles of speaker and listener switch back and forth in an exchange, and usually involve covert speaker and listener behavior as well.

verbal function-altering effect Verbal stimuli can alter the functional effects of immediate or future SD s and MOs and, accordingly, change a listener's behavior. For example, being told "The bridge is out, turn left at the 7-Eleven and there will be another one in 5 miles" can alter the functional effects of stimuli encountered in the future and evoke verbal and nonverbal behavior at that time (e.g., tacting the 7-Eleven, turning left).

verification One of three components of the experimental reasoning, or baseline logic, used in single-subject research designs; accomplished by demonstrating that the prior level of baseline responding would have remained unchanged had the independent variable not been introduced. Verifying the accuracy of the original prediction reduces the probability that some uncontrolled (confounding) variable was responsible for the observed change in behavior. (See also **prediction** and **replication.**)

video modeling A behavior change strategy in which the participant views a video of a model performing the target behavior and then imitates the behavior.

video self-modeling A behavior change strategy in which the participant views a video of himself successfully performing the target behavior and then imitates his own model.

visual analysis A systematic approach for interpreting the results of behavioral research and treatment programs that entails visual inspection of graphed data for level, trend, and variability within and between experimental conditions.

W

whole-interval recording A time sampling method for measuring behavior in which the observation period is divided into a series of brief time intervals (typically from 5 to 15 seconds). At the end of each interval, the observer records whether the target behavior occurred throughout the entire interval; tends to underestimate the overall percentage of the observation period in which the target behavior actually occurred.

withdrawal design A term used by some authors as a synonym for A-B-A-B design; also used to describe experiments in which an effective treatment is sequentially or partially withdrawn to promote the maintenance of behavior changes. (See also **A-B-A-B design** and **reversal design.**)

BIBLIOGRAPHY

Abernathy, W. B., & Lattal, D. (2014). Organizational behavior management: Past, present, and future. In F. K. McSweeney & E. S. Murphy (Eds.), *The Wiley-Blackwell handbook of operant and classical conditioning* (pp. 645–668). West Sussex, UK: John Wiley & Sons.

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology*, *87*, 49–74.

Adams, B. J., Fields, L., & Verhave, T. (1993). Effects of test order on intersubject variability during equivalence class formation. *The Psychological Record*, *43*, 133–152.

Adams, C., & Kelley, M. (1992). Managing sibling aggression: Overcorrection as an alternative to timeout. *Behavior Modification, 23,* 707–717.

Adams, K. A., & Lawrence, E. K. (2019). *Research methods, statistics, and applications* (2nd ed.). Thousand Oaks, CA: Sage.

Addison, L., & Lerman, D. C. (2009). Descriptive analysis of teachers' responses to problem behavior following training. *Journal of Applied Behavior Analysis*, *42*, 485–490.

Adelinis, J. D., Piazza, C. C., & Goh, H. L. (2001). Treatment of multiply controlled destructive behavior with food reinforcement. *Journal of Applied Behavior Analysis*, *34*, 97–100.

Adkins, V. K., & Mathews, R. M. (1997). Prompted voiding to reduce incontinence in community-dwelling older adults. *Journal of Applied Behavior Analysis*, *30*, 153–156.

Aguirre A. A., Valentino, A. L., & LeBlanc, L. A. (2016). Empirical investigations of the intraverbal since 2005. *The Analysis of Verbal Behavior*, *32*, 139–153.

Ahearn, W. H. (2003). Using simultaneous presentation to increase vegetable consumption in a mildly selective child with autism. *Journal of Applied Behavior Analysis*, *36*, 361–365.

Ahearn, W. H., & Tiger, J. H. (2013). Behavioral approaches to the treatment of autism. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. 301–327). Washington, DC: American Psychological Association.

Ahearn, W. H., Clark, K. M., Gardenier, N. C., Chung, B. I., & Dube, W. V. (2003). Persistence of stereotypic behavior: Examining the effects of external reinforcers. *Journal of Applied Behavior Analysis*, *36*, 439–448.

Ahearn, W. H., Clark, K. M., MacDonald, R. P. F., & Chung, B. I. (2007). Assessing and treating vocal stereotypy in children with autism. *Journal of Applied Behavior Analysis*, *40*, 263–275.

Ahearn, W. H., Kerwin, M. E., Eicher, P. S., Shantz, J., & Swearingin, W. (1996). An alternating treatments comparison of two intensive interventions for food refusal. *Journal of Applied Behavior Analysis*, *29*, 321–332.

Ahrens, E. N., Lerman, D. C., Kodak, T., Worsdell, A. S., & Keegan, C. (2011). Further evaluation of response interruption and redirection as treatment for stereotypy. *Journal of Applied Behavior Analysis*, *44*, 95–108.

Ainslie, G. (1974). Impulse control in pigeons. *Journal of the Experimental Analysis of Behavior*, *21*, 485–489.

Alber, S. R., & Heward, W. L. (1997). Recruit it or lose it! Training students to recruit contingent teacher attention. *Intervention in School and Clinic*, *5*, 275–282.

Alber, S. R., & Heward, W. L. (2000). Teaching students to recruit positive attention: A review and recommendations. *Journal of Behavioral Education*, *10*, 177–204.

Alber, S. R., Anderson, L. L., Martin, C. M., & Moore, K. J. (2004). Teaching elementary students with behavior disorders to recruit positive teacher attention: Effects on math proficiency. *Journal of Early and Intensive Behavior Intervention*, *1*, 218–231.

Alber, S. R., Heward, W. L., & Hippler, B. J. (1999). Training middle school students with learning disabilities to recruit positive teacher attention. *Exceptional Children*, *65*, 253–270.

Alber, S. R., Nelson, J. S., & Brennan, K. B. (2002). A comparative analysis of two homework study methods on elementary and secondary students' acquisition and maintenance of social studies content. *Education and Treatment of Children*, *25*, 172–196.

Albert, K. M., Carbone, V. J., Murray, D. D., Hagerty, M., & Sweeney-Kerwin, E. J. (2012). Increasing the mand repertoire of children with autism through the use of an interrupted chain procedure. *Behavior Analysis in Practice*, *5*, 65–76.

Alberto, P. A., & Troutman, A. C. (1982). *Applied behavior analysis for teachers*. Columbus, OH: Merrill.

Alberto, P. A., & Troutman, A. C. (2013). *Applied behavior analysis for teachers* (9th ed.). Upper Saddle River, NJ: Pearson.

Albin, R. W., & Horner, R. H. (1988). Generalization with precision. In R. H. Horner, G. Dunlap, & R. L. Koegel (Eds.), *Generalization and maintenance: Life-style changes in applied settings* (pp. 99–120). Baltimore: Brookes.

Albright, L., Reeve, K. F., Reeve, S. A., & Kisamore, A. (2015). Teaching statistical variability with equivalence-based instruction. *Journal of Applied Behavior Analysis*, *48*, 883–894.

Aldi, C., Crigler, A., Kates-McElrath, K., Long, B., Smith, H., Rehak, K., & Wilkinson, L. (2016). Examining the effects of video modeling and prompts to teach activities of daily living skills. *Behavior Analysis in Practice*, *9*, 384–388. doi 10.1007/s40617-016-0127-y

Alessi, G. (1987). Generative strategies and teaching for generalization. *The Analysis of Verbal Behavior*, *5*(1), 15–27.

Alessi, G. (1992). Models of proximate and ultimate causation in psychology. *American Psychologist*, *48*, 1359–1370.

Allan, A. C., Vladescu, J. C., Kisamore, A. N., Reeve, S. A., & Sidener, T. M. (2014). Evaluating the emergence of reverse intraverbals in children with autism. *The Analysis of Verbal Behavior*, *30*, 1–17. doi:10.1007/s40616-014-0025-8

Allen, D. (2001). *Getting things done: The art of stress-free productivity*. New York: Viking.

Allen, J. S., Jr., Tarnowski, K. J., Simonian, S. J., Elliot, D. & Drabman, R. S. (1991). The generalization map revisited: Assessment of generalized treatment effects in child and adolescent behavior therapy. *Behavior Therapy*, *22*, 393–405.

Allen, K. D., & Evans, J. H. (2001). Exposure-based treatment to control excessive blood glucose monitoring. *Journal of Applied Behavior Analysis*, *34*, 497–500.

Allen, K. D., Wallace, D. P., Renes, D., Bowen, S. L., & Burke, R. V. (2010). Use of video modeling to teach vocational skills to adolescents and young adults with autism spectrum disorders. *Education and Treatment of Children*, *33*(3), 339–349.

Allen, K. E., Hart, B. M., Buell, J. S., Harris, F. R., & Wolf, M. M. (1964). Effects of social reinforcement on isolate behavior of a nursery school child. *Child Development*, *35*, 511–518.

Allen, L. D., Gottselig, M., & Boylan, S. (1982). A practical mechanism for using free time as a reinforcer in the classroom. *Education and Treatment of Children*, *5*(4), 347–353.

Alling, K., & Poling, A. (1995). The effects of differing response-force requirements on fixed-force requirements on fixed-ratio responding of rats. *Journal of the Experimental Analysis of Behavior*, *63*(3), 331–346.

Allison, J. (1993). Response deprivation, reinforcement, and economics. *Journal of the Experimental Analysis of Behavior*, *60*, 129–140.

Alloway, T., Wilson, G., & Graham, J. (2005). Sniffy the virtual rat, Version 3.0. Belmont, CA: Wadsworth.

Almeida-Verdu, A. C., Huziwara, E. M., de Souza, D. G., de Rose, J. C., Bevilacqua, M. C., Lopes J., Jr., . . . McIlvane, W. J. (2008). Relational learning in children with deafness and cochlear implants. *Journal of the Experimental Analysis of Behavior*, *89*, 407–424.

Alstot, A. E. (2015). Effects of a peer-administered token economy on second grade physical education students' overhand throw performance. *Physical Educator*, *72*(2), 245–265.

Alter, M. M., & Borrero, J. C. (2015). Teaching generatively: Learning about disorders and disabilities. *Journal of Applied Behavior Analysis*, *48*, 376–389.

Alter, P. J., Conroy, M. A., Mancil, G. R., & Haydon, T. (2008). A comparison of functional behavior assessment methodologies with young children: Descriptive methods and functional analysis. *Journal of Behavioral Education*, *17*, 200–219. doi: http://dx.doi.org.libproxy.library.wmich.edu/10.1007/s10864-008-9064-3

Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behaviour*, *49*(3/4), 227–267.

Altschuld, J. W., & Witkin, B. R. (2000). *From needs assessment to action: Transforming needs into solution strategies*. Thousand Oaks, CA: Sage Publications, Inc.

Altus, D. E., & Morris, E. K. (2009). B. F. Skinner's utopian vision: Behind and beyond Walden Two. *The Behavior Analyst*, *32*, 319–335.

Alvero, A. M., Struss, K., & Rappaport, E. (2007). Measuring safety performance: A comparison of whole, partial, and momentary time-sampling recording methods. *Journal of Organizational Behavior Management*, *27*, 1–28.

Alwell, M., Hunt, P., Goetz, L., & Sailer, W. (1989). Teaching generalized communicative behaviors within interrupted behavior chain contexts. *The Journal of the Association for Persons with Severe Handicaps*, *14*, 91–101.

Amd, M., & Roche, B. (2018). Assessing the effects of a relational skills training intervention on fluid intelligence among a sample of socially-disadvantaged children in Bangladesh. *The Psychological Record*, *68*, 141–149.

American Football Coaches Association. (1995). Football coaching strategies. Champaign, IL: Human Kinetics.

American Psychological Association. (2010). *Ethical principles of psychologists and code of conduct*. Washington, DC: Author. Retrieved February 1, 2017, from http://www.apa.org/ethics/code/

American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

American Psychological Association. Ethics Code. (2004). http://www.apa.org/ethics/code/index.aspx. Retrieved June 1, 2017.

Anderson, C. M., & Long, E. S. (2002). Use of a structured descriptive assessment methodology to identify variables affecting problem behavior. *Journal of Applied Behavior Analysis*, *35*, 137–154.

Anderson, M. T., Vu, C., Derby, K. M., Goris, M., & McLaughlin, T. F. (2002). Using functional analysis procedures to monitor medication effects in an outpatient and school setting. *Psychology in the Schools*, *39*, 73–76.

Andics, A., Gabor, A., Gacsi, M., Farago, T., Szabo, D., & Miklosi, A. (2016). Neural mechanisms for lexical processing in dogs. *Science*, *353*, 1030–1032.

Andresen, J. T. (1991). Skinner and Chomsky 30 years later. Or: The return of the repressed. *The Behavior Analyst*, *14*, 49–60.

Andronis, P. T. (1983). *Symbolic aggression by pigeons: Contingency coadduction*. Unpublished doctoral dissertation, University of Chicago, Department of Psychiatry and Behavior Analysis, Chicago.

Andronis, P. T., Layng, T. V. J., & Goldiamond, I. (1997). Contingency adduction of "symbolic aggression" by pigeons. *The Analysis of Verbal Behavior*, *14*, 5–17.

Andzik, N. R., Cannella-Malone, H. I., & Sigafoos, J. (2016). Practitioner implemented functional communication training: A review of the literature. *Research and Practice for Persons with Severe Disabilities*, *41*, 79–89. doi:10.1177/1540796916633874

Annett, J. M., & Leslie, J. C. (1995). Stimulus equivalence classes involving olfactory stimuli. *The Psychological Record*, *45*, 439–450.

Ardoin, S. P., Martens, B. K., & Wolfe, L. A. (1999). Using high-probability instruction sequences with fading to increase student compliance during transitions. *Journal of Applied Behavior Analysis*, *32*, 339–351.

Armendariz, F., & Umbreit, J. (1999). Using active responding to reduce disruptive behavior in a general education classroom. *Journal of Positive Behavior Interventions*, *1*, 152–158.

Arndorfer, R., & Miltenberger, R. (1993). Functional assessment and treatment of challenging behavior: A review with implications for early childhood. *Topics in Early Childhood Special Education*, *13*, 82–105.

Arntzen, E. (2012). Training and testing parameters in formation of stimulus equivalence: Methodological issues. *European Journal of Behavior Analysis*, *13*, 123–135.

Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record*, *47*, 309–320.

Arntzen, E., Halstadtrø, A., & Halstadtrø, M. (2003). Training play behavior in a 5-year-old boy with developmental disabilities. *Journal of Applied Behavior Analysis*, *36*, 367–370.

Arntzen, E., Halstadtrø, L., Bjerke, E., & Halstadtrø, M. (2010). Training and testing music skills in a boy with autism using a matching-to-sample format. *Behavioral Interventions*, *25*, 129–143.

Arsenault, D. J., Smith, L. D., & Beauchamp, E. A. (2006). Visual inscriptions in scientific hierarchy: Mapping the "treasures of science." *Science Communication*, *27*, 376–428.

Artman, K., Wolery, M., & Yoder, P. (2012). Embracing our visual inspection and analysis tradition: Graphing interobserver agreement data. *Remedial and Special Education*, *33*(2), 71–77. doi:10.1177/0741932510381653.

Ashbaugh, R., & Peck, S. M. (1998). Treatment of sleep problems in a toddler: A replication of the faded bedtime with response cost protocol. *Journal of Applied Behavior Analysis*, *31*, 127–129.

Aspen, A., & Stack. L. (2017). *Visual activity schedules for students with autism spectrum disorder*. California Autism Professional Training and Information Network. https://www.youtube.com/watch?v=YNfnuuATlkA

Assaz, D. A., Roche, B., Kanter, J. W., & Oshiro, C. K. B. (2018). Cognitive defusion in Acceptance and Commitment Therapy: What are the basic processes of change? *Psychological Record*, *68*, 405–418.

Association for Behavior Analysis International. (1989). *Statement on the right to effective behavioral treatment, 1989*. Kalamazoo, MI: Author. Retrieved September 29, 2016, from https://www.abainternational.org/about-us/policies-and-positions/right-to-effective-behavioral-treatment,-1989.aspx

Association for Behavior Analysis International. (1990). *Statement on students' right to effective education, 1990*. Kalamazoo, MI: Author. Retrieved November 11, 2006, from https://www.abainternational.org/about-us/policies-and-positions/students-rights-to-effective-education,-1990.aspx

Association for Behavior Analysis International. (2010). *Statement on restraint and seclusion, 2010*. Kalamazoo, MI: Author. Retrieved September 29, 2016, from https://www.abainternational.org/about-us/policies-and-positions/restraint-and-seclusion,-2010.aspx

Association for Behavior Analysis International (ABAI). (2012). *Behavior Change for a Sustainable World Conference*. Program retrieved August 21, 2018, at https://www.abainternational.org/events/program-details/event-detail.aspx?intConvId=25&by=Day&date=8/3/2012

Association for Persons with Severe Handicaps. (1987, May). Resolution on the cessation of intrusive interventions. *TASH Newsletter*, *5*, 3.

Athens, E. S., & Vollmer, T. R. (2010). An investigation of differential reinforcement of alternative behavior without extinction. *Journal of Applied Behavior Analysis*, *43*, 569–589. http://dx.doi.org/10.1901/jaba.2010.43-569

Athens, E. S., Vollmer, T. R., & St. Peter Pipkin, C. C. (2007). Shaping academic task engagement with percentile schedules. *Journal of Applied Behavior Analysis*, *40*, 475–488.

Athens, E. S., Vollmer, T. R., Sloman, K. N., & St. Peter Pipkin, C. (2008). An analysis of vocal stereotypy and therapist fading. *Journal of Applied Behavior Analysis*, *41*, 291–297.

A-Tjak, J. G., Davis, M. L., Morina, N., Powers, M. B., Smits, J. A., & Emmelkamp, P. M. (2015). A meta-analysis of the efficacy of acceptance and commitment therapy for clinically relevant mental and physical health problems. *Psychotherapy and Psychosomatics*, *84*, 30–36.

Atwater, J. B., & Morris, E. K. (1988). Teachers' instructions and children's compliance in preschool classrooms: A descriptive analysis. *Journal of Applied Behavior Analysis*, *21*, 157–167.

Audubon, J. J. (1827–1838). *The birds of America*. London, UK: Author.

Augustson, E. M., & Dougher, M. J. (1997). The transfer of avoidance evoking functions through stimulus equivalence classes. *Journal of Behavior Therapy and Experimental Psychiatry*, *28*, 181–191.

Augustson, E. M., Dougher, M. J., & Markham, M. R. (2000). Emergence of conditional stimulus relations and transfer of respondent eliciting functions among compound stimuli. *Journal of the Experimental Analysis of Behavior*, *50*, 745–770.

Augustson, K. G., & Dougher, M. J. (1992). Teaching conditional discrimination to young children: Some methodological successes and failures. *Experimental Analysis of Human Behavior Bulletin*, *9*, 21–24.

Austin, J. L., & Bevan, D. (2011). Using differential reinforcement of low rates to reduce children's requests for teacher attention. *Journal of Applied Behavior Analysis*, *44*, 451–461. doi:10.1901/jaba.2011.44-451

Austin, J. L., Groves, E. A., Reynish, L. C., & Francis, L. L. (2015). Validating trial-based functional analyses in mainstream primary school classrooms. *Journal of Applied Behavior Analysis*, *48*, 274–288. doi.org/10.1002/jaba.208

Axe, J. B. (2008). Conditional discrimination in the intraverbal relation: A review and recommendations for future research. *The Analysis of Verbal Behavior*, *24*, 159–174.

Axe, J. B., & Laprime, A. P. (2017). The effect of contingent pairing on establishing praise as a reinforcer with children with autism. *Journal*

of Developmental and Physical Disabilities, *29*, 325–340. doi:10.1007/s10882-016-9526-9

Axe, J. B., Phelan, S. H., & Irwin, C. L. (in press). A review and discussion of applied behavior analytic research on problem solving. *The Analysis of Verbal Behavior*.

Axe, J. B., & Sainato, D. M. (2010). Matrix training of preliteracy skills with preschoolers with autism. *Journal of Applied Behavior Analysis*, *43*, 635–652. doi:10.1901/jaba.2010.43-635

Axelrod, M. J., & Zank, A. J. (2012). Increasing classroom compliance: Using a high-probability command sequence with noncompliant students. *Journal of Behavioral Education*, *21*, 119–133. doi:11.1007/s10864-011-9145-6

Axelrod, S. (2017). A commentary on Critchfield and Reed: The fuzzy concept of applied behavior analysis research. *The Behavior Analyst*, *40*, 167–171. doi 10.1007/s40614-017-0117-6

Axelrod, S. A. (1990). Myths that (mis)guide our profession. In A. C. Repp & N. N. Singh (Eds.), *Perspectives on the use of nonaversive and aversive interventions for persons with developmental disabilities* (pp. 59–72). Sycamore, IL: Sycamore.

Axelrod, S., Hall, R. V., Weis, L., & Rohrer, S. (1971). *Use of self-imposed contingencies to reduce the frequency of smoking behavior*. Paper presented at the Fifth Annual Meeting of the Association for the Advancement of Behavior Therapy, Washington, DC.

Axelrod, S., Spreat, S., Berry, B., & Moyer, L. (1993). A decision-making model for selecting the optimal treatment procedures. In R. Van Houten and S. Axelrod (Eds.), *Behavior analysis and treatment* (pp. 183–202). New York: Plenum Press.

Ayllon, T., & Azrin, N. H. (1968). *The token economy: A motivational system for therapy and rehabilitation*. New York: Appleton-Century-Crofts.

Ayllon, T., & Michael, J. (1959). The psychiatric nurse as a behavioral engineer. *Journal of the Experimental Analysis of Behavior*, *2*, 323–334.

Azrin, N. H. (1960). Effects of punishment intensity during variable-interval reinforcement. *Journal of Experimental Analysis of Behavior*, *3*, 123–142.

Azrin, N. H. (1960). Sequential effects of punishment. *Science*, *131*, 605–606.

Azrin, N. H., & Besalel, V. A. (1980). *How to use overcorrection*. Austin, TX: PRO-ED.

Azrin, N. H., & Besalel, V. A. (1999). *How to use positive practice, self-correction, and overcorrection* (2nd ed.). Austin, TX: PRO-ED.

Azrin, N. H., & Foxx, R. M. (1971). A rapid method of toilet training the institutionalized retarded. *Journal of Applied Behavior Analysis*, *4*, 89–99.

Azrin, N. H., & Holz, W. C. (1966). Punishment. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 380–447). New York: Appleton-Century-Crofts.

Azrin, N. H., & Nunn, R. G. (1973). Habit-reversal for habits and tics. *Behavior Research and Therapy*, *11*, 619–628.

Azrin, N. H., & Peterson, A. L. (1988). Habit reversal for the treatment of Tourette syndrome. *Behaviour Research and Therapy*, *26*, 347–351. doi:10.1016/0005-7967(88)90089-7

Azrin, N. H., & Powers, M. A. (1975). Eliminating classroom disturbances of emotionally disturbed children by positive practice procedures. *Behavior Therapy*, *6*, 525–534.

Azrin, N. H., & Wesolowski, M. D. (1974). Theft reversal: An overcorrection procedure for eliminating stealing by retarded persons. *Journal of Applied Behavior Analysis*, *7*, 577–581.

Azrin, N. H., Flores, T., & Kaplan, S. J. (1975). Job-finding club: A group-assisted program for obtaining employment. *Behaviour Research & Therapy*, *13*, 17–27.

Azrin, N. H., Holz, W. C., & Hake, D. C. (1963). Fixed-ratio punishment by intense noise. *Journal of the Experimental Analysis of Behavior*, *6*, 141–148.

Azrin, N. H., Hutchinson, R. R., & Hake, D. C. (1963). Pain-induced fighting in the squirrel monkey. *Journal of the Experimental Analysis of Behavior*, *6*, 620.

Azrin, N. H., Kaplan, S. J., & Foxx, R. M. (1973). Autism reversal: Eliminating stereotyped self stimulation of retarded individuals. *American Journal of Mental Deficiency*, *78*, 241–248.

Azrin, N. H., Nunn, R. G., & Frantz, S. E. (1980). Habit reversal vs. negative practice treatment of nail biting. *Behavior Research and Therapy*, *18*, 281–285.

Azrin, N. H., Rubin, H., O'Brien, F., Ayllon, T., & Roll, D. (1968). Behavioral engineering: Postural control by a portable operant apparatus. *Journal of Applied Behavior Analysis*, *1*, 99–108.

Babcock, R. A., Sulzer-Azaroff, B., Sanderson, M., & Scibak, J. (1992). Increasing nurses' use of feedback to promote infection-control practices in a head-injury treatment center. *Journal of Applied Behavior Analysis*, *25*, 621–627.

Babyak, A. E., Luze, G. J., & Kamps, D. M. (2000). The Good Student Game: Behavior management for diverse classrooms. *Intervention in School and Clinic*, *35*(4), 216–223.

Baer, D. M. (1960). Escape and avoidance response of preschool children to two schedules of reinforcement withdrawal. *Journal of the Experimental Analysis of Behavior*, *3*, 155–159.

Baer, D. M. (1961). Effect of withdrawal of positive reinforcement on an extinguishing response in young children. *Child Development*, *32*, 67–74.

Baer, D. M. (1962). Laboratory control of thumbsucking by withdrawal and representation of reinforcement. *Journal of the Experimental Analysis of Behavior*, *5*, 525–528.

Baer, D. M. (1971). Let's take another look at punishment. *Psychology Today*, *5*, 5–32.

Baer, D. M. (1975). In the beginning, there was the response. In E. Ramp & G. Semb (Eds.), *Behavior analysis: Areas of research and application* (pp. 16–30). Upper Saddle River, NJ: Prentice Hall.

Baer, D. M. (1977a). Reviewer's comment: Just because it's reliable doesn't mean that you can use it. *Journal of Applied Behavior Analysis*, *10*, 117–119.

Baer, D. M. (1977b). "Perhaps it would be better not to know everything." *Journal of Applied Behavior Analysis*, *10*, 167–172.

Baer, D. M. (1981). A hung jury and a Scottish verdict: "Not proven." *Analysis and Intervention in Developmental Disabilities*, *1*, 91–97.

Baer, D. M. (1982). Applied behavior analysis. In G. T. Wilson & C. M. Franks (Eds.), *Contemporary behavior therapy: Conceptual and empirical foundations* (pp. 277–309). New York: Guilford Press.

Baer, D. M. (1985). [Symposium discussant]. In C. E. Naumann (Chair), *Developing response classes: Why re-invent the wheel?* Symposium conducted at the Annual Conference of the Association for Behavior Analysis, Columbus, OH.

Baer, D. M. (1987). Weak contingencies, strong contingencies, and too many behaviors to change. *Journal of Applied Behavior Analysis*, *20*, 335–337.

Baer, D. M. (1991). Tacting "to a fault." *Journal of Applied Behavior Analysis*, *24*, 429–431.

Baer, D. M. (1999). *How to plan for generalization* (2nd ed.). Austin, TX: PRO-ED.

Baer, D. M. (2005). Letters to a lawyer. In W. L. Heward, T. E. Heron, N. A. Neef, S. M. Peterson, D. M. Sainato, G. Cartledge, R. Gardner, III, L. Peterson, S.B. Hersh, & J. C. Dardig (Eds.), *Focus on behavior analysis in education: Achievements, challenges and opportunities* (pp. 5–30). Upper Saddle River, NJ: Prentice Hall.

Baer, D. M., & Fowler, S. A. (1984). How should we measure the potential of self-control procedures for generalized educational outcomes? In W. L. Heward, T. E. Heron, D. S. Hill, & J. Trap-Porter (Eds.), *Focus on behavior analysis in education* (pp. 145–161). Columbus, OH: Charles E. Merrill.

Baer, D. M., & Sherman, J. A. (1964). Reinforcement control of generalized imitation in young children. *Journal of Experimental Child Psychology*, *1*, 37–49.

Baer, D. M., & Wolf, M. M. (1970a). Recent examples of behavior modification in preschool settings. In C. Neuringer & J. L. Michael (Eds.), *Behavior modification in clinical psychology* (pp. 10–55). Upper Saddle River, NJ: Prentice Hall.

Baer, D. M., & Wolf, M. M. (1970b). The entry into natural communities of reinforcement. In R. Ulrich, T. Stachnik, & J. Mabry (Eds.), *Control of human behavior* (Vol. 2, pp. 319–324). Glenview, IL: Scott, Foresman.

Baer, D. M., Peterson, R. F., & Sherman, J. A. (1967). The development of imitation by reinforcing behavioral similarity of a model. *Journal of the Experimental Analysis of Behavior*, *10*, 405–416.

Baer, D. M., Wolf, M. M., & Risley, T. (1968). Current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*, 91–97.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1987). Some still-current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *20*, 313–327.

Baer, R. A. (2003). Mindfulness training as a clinical intervention: A conceptual and empirical review. *Clinical Psychology: Science and Practice*, *10*, 125–143.

Baer, R. A., Blount, R. L., Detrich, R., & Stokes, T. F. (1987). Using intermittent reinforcement to program maintenance of verbal/nonverbal correspondence. *Journal of Applied Behavior Analysis*, *20*, 179–184.

Baer, R. A., Tishelman, A. C., Degler, J. D., Osnes, P. G., & Stokes, T. F. (1992). Effects of self vs. experimenter-selection of rewards on classroom behavior in young children. *Education and Treatment of Children*, *15*, 1–14.

Baer, R. A., Williams, J. A., Osnes, P. G., & Stokes, T. F. (1984). Delayed reinforcement as an indiscriminable contingency in verbal/nonverbal correspondence training. *Journal of Applied Behavior Analysis*, *17*, 429–440.

Bailey, D. B., Jr., & Wolery, M. (1992). *Teaching infants and preschoolers with disabilities* (2nd ed). Upper Saddle River, NJ: Merrill/Prentice Hall.

Bailey, J., & Burch, M. (2010). *25 Essential skills and strategies for professional behavior analysts*. New York: Routledge.

Bailey, J., & Burch, M. (2016). *Ethics for behavior analysts* (3rd ed.). New York: Routledge.

Bailey, J., & Meyerson, L. (1969). Vibration as a reinforcer with a profoundly retarded child. *Journal of Applied Behavior Analysis*, *2*, 135–137.

Bailey, J. S., & Pyles, D. A. M. (1989). Behavioral diagnostics. In E. Cipani (Ed.), *The treatment of severe behavior disorders: Behavior analysis approach* (pp. 85–107). Washington, DC: American Association on Mental Retardation.

Bailey, S. L., & Lessen, E. I. (1984). An analysis of target behaviors in education: Applied but how useful? In W. L. Heward, T. E. Heron, D. S. Hill, & J. Trap-Porter (Eds.), *Focus on behavior analysis in education* (pp. 162–176). Columbus, OH: Charles E. Merrill.

Baker, J. C., & LeBlanc, L. A. (2014). Aging. In F. K. McSweeney & E. S. Murphy (Eds.), *The Wiley-Blackwell handbook of operant and classical conditioning* (pp. 695–713). Oxford, UK: John Wiley & Sons.

Baker, J. C., Hanley, G. P., & Mathews, R. M. (2006). Staff-administered functional analysis and treatment of aggression by an elder with dementia. *Journal of Applied Behavior Analysis*, *39*, 469–474.

Baker, J. C., LeBlanc, L. A., & Raetz, P. B. (2008). A behavioral conceptualization of aphasia. *The Analysis of Verbal Behavior*, *24*, 147–158.

Ballard, K. D., & Glynn, T. (1975). Behavioral self-management in story writing with elementary school children. *Journal of Applied Behavior Analysis*, *8*, 387–398.

Bancroft, S. L., Weiss, J. S., Libby, M. E., & Ahearn, W. H. (2011). A comparison of procedural variations in teaching behavior chains: Manual guidance, trainer completion, and no completion of untrained steps. *Journal of Applied Behavior Analysis*, *44*, 559–569.

Banda, D. R., Grimmett, E., & Hart, S. L. (2009). Activity schedules: Helping students with autism spectrum disorders in general education classrooms manage transition issues. *Teaching Exceptional Children*, *41*(4), 16–21.

Bandura, A. (1969). *Principles of behavior modification*. New York: Holt, Rinehart & Winston.

Bandura, A. (1971). Vicarious and self-reinforcement processes. In R. Glaser (Ed.), *The nature of reinforcement*. New York: Academic Press.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A., Ross, D., & Ross, S. A. (1963). Vicarious reinforcement and imitative learning. *Journal of Abnormal and Social Psychology*, *67*, 601–607. doi:10.1037/h0045550

Bannerman, D. J., Sheldon, J. B., Sherman, J. A., & Harchik, A. E. (1990). Balancing the right to habilitation with the right to personal liberties: The rights of people with developmental disabilities to eat too many doughnuts and take a nap. *Journal of Applied Behavior Analysis*, *23*, 79–89. doi:10.1901/jaba.1990.23-79

Barbera, M. L., & Rasmussen, T. (2007). *The verbal behavior approach*. London, UK: Jessica Kingsley Publishers.

Barbetta, P. M., Heron, T. E., & Heward, W. L. (1993). Effects of active student response during error correction on the acquisition, maintenance, and generalization of sight words by students with developmental disabilities. *Journal of Applied Behavior Analysis*, *26*, 111–119.

Barker, M. R., Bailey, J. S., & Lee, N. (2004). The impact of verbal prompts on child safety-belt use in shopping carts. *Journal of Applied Behavior Analysis*, *37*, 527–530.

Barlow, D. H., & Hayes, S. C. (1979). Alternating treatments design: One strategy for comparing the effects of two treatments in a single behavior. *Journal of Applied Behavior Analysis*, *12*, 199–210.

Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior change*. Boston: Pearson Education.

Barnes-Holmes, D. (2000). Behavioral pragmatism: No place for reality and truth. *The Behavior Analyst*, *23*, 191–202.

Barnes-Holmes, D., & Barnes-Holmes, Y. (2000). Explaining complex behavior: Two perspectives on the concept of generalized operant classes. *The Psychological Record*, *50*, 251–265.

Barnes-Holmes, D., Hayes, S. C., & Dymond, S. (2001). Self and self-directed rules. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 119–139). New York: Plenum.

Barnes-Holmes, D., Hayes, S. C., Dymond, S., & O'Hora, D. (2001). Multiple stimulus relations and the transformation of stimulus functions. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 51–71). New York: Plenum.

Barnes-Holmes, D., Hayes, S. C., & Gregg, J. (2001). Religion, spirituality, and transcendence. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 229–251). New York: Plenum.

Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., Healy, O., Lyddy, F., Cullinan, V., & Hayes, S. C. (2001). Psychological development. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 157–180). New York: Plenum.

Barnes-Holmes, Y., Barnes-Holmes, D., Smeets, P. M., Strand, P., & Friman, P. (2004). Establishing relational responding in accordance with more-than and less-than as generalized operant behavior in young children. *International Journal of Psychology and Psychological Therapy*, *4*, 531–558.

Barnes-Holmes, Y., McHugh, L., & Barnes-Holmes, D. (2004). Perspective-talking and theory of mind: A relational frame account. *The Behavior Analyst Today*, *5*, 15–24.

Barnoy, E. L., Najdowski, A. C., Tarbox, J., Wilke, A. E., Nollet, M. D., & Roane, H. (2009). Evaluation of a multicomponent intervention for diurnal bruxism in a young child with autism. *Journal of Applied Behavior Analysis*, *42*, 845–848.

Bar-On, R. E., & Parker, J. D. (2000). *The handbook of emotional intelligence: Theory, development, assessment, and application at home, school, and in the workplace*. San Francisco: Jossey-Bass.

Baron, A., & Galizio, M. (2005). Positive and negative reinforcement: Should the distinction be preserved? *The Behavior Analyst*, *28*, 85–98.

Baron, A., & Galizio, M. (2006). The distinction between positive and negative reinforcement: Use with care. *The Behavior Analyst*, *29*, 141–151.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*, 37–46.

Barretto, A., Wacker, D. P., Harding, J., Lee, J., & Berg, W. K. (2006). Using telemedicine to conduct behavioral assessments. *Journal of Applied Behavior Analysis*, *39*, 333–340.

Barrish, H. H., Saunders, M., & Wolf, M. M. (1969). Good behavior game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis*, *2*, 119–124.

Barron, B. F., Verkuylen, L., Belisle, J., Paliliunas, D., & Dixon, M. R. (2018). Teaching "then-later" and "here-there" relations to children with autism: An evaluation of single reversals and transformation of stimulus function. *Behavior Analysis in Practice*. Advance online publication.

Barron, R., Leslie, J. C., & Smyth, S. (2018). Teaching real-world categories using touchscreen equivalence-based instruction. *The Psychological Record*, *68*, 89–101.

Barry, A. K. (1998). *English grammar: Language as human behavior*. Upper Saddle River, NJ: Prentice Hall.

Barry, L. M., & Messer, J. J. (2003). A practical application of self-management for students diagnosed with attention-deficit/hyperactivity disorder. *Journal of Positive Behavior Interventions*, *5*, 238–248.

Barton, E. J., & Osborne, J. G. (1978). The development of classroom sharing by a teacher using positive practice. *Behavior Modification*, *2*(2), 231–250. https://doi.org/10.1177/014544557822005

Barton, E. S., Guess, D., Garcia, E., & Baer, D. M. (1970). Improvement of retardates' mealtime behaviors by timeout procedures using multiple baseline techniques. *Journal of Applied Behavior Analysis*, *3*, 77–84.

Barton, L. E., Brulle, A. R., & Repp, A. C. (1986). Maintenance of therapeutic change by momentary DRO. *Journal of Applied Behavior Analysis*, *19*, 277–282.

Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective taking: Imagining how another feels versus imaging how you would feel. *Personality and Social Psychology Bulletin, 23*, 751–758.

Baum, W. M. (1994). *Understanding behaviorism: Science, behavior, and culture.* New York: Harper Collins.

Baum, W. M. (2011). Behaviorism, private events, and the molar view of behavior. *The Behavior Analyst, 34*, 185–200.

Baum, W. M. (2011). What is radical behaviorism? A review of Jay Moore's Conceptual Foundations of Radical Behaviorism. *Journal of the Experimental Analysis of Behavior, 95*, 119–126.

Baum, W. M. (2017). *Understanding behaviorism: Behavior, culture, and evolution* (3rd ed.). Chichester, UK: John Wiley & Sons.

Bayley, N. (2005). *Bayley Scales of Infant and Toddler Development* (3rd ed.). San Antonio, TX: PsychCorp.

Beavers, G. A., Iwata, B. A., & Gregory, M. K. (2014). Parameters of reinforcement and respose-class hierarchies. *Journal of Applied Behavior Analysis, 47*, 70–82.

Beavers, G. A., Iwata, B. A., & Lerman, D. C. (2013). Thirty years of research on the functional analysis of problem behavior. *Journal of Applied Behavior Analysis, 46*, 1–21.

Becker, W. C. (1986). *Applied psychology for teachers: A cognitive behavioral approach.* Chicago: Science Research Associates.

Becker, W. C., & Engelmann, S. E. (1978). Systems for basic instruction: Theory and applications. In A. Catania & T. Brigham (Eds.), *Handbook of applied behavior analysis: Social and instructional processes.* New York: Irvington.

Becker, W. C., Engelmann, S., & Thomas, D. R. (1975). *Teaching 2: Cognitive learning and instruction.* Chicago: Science Research Associates.

Becraft, J. L., Borrero, J. C., Davis, B. J., & Mendres-Smith, A. E. (2016). Assessment of a rotating time sampling procedure: Implications for interobserver agreement and response measurement. *Education and Treatment of Children, 39*, 1–20.

Behavior Analyst Certification Board (BACB). (2014). *Professional and ethical compliance code for behavior analysts.* Retrieved June 22, 2017, from www.bacb.com/ethics-code/

Behavior Analyst Certification Board (BACB). (2017). *BCBA/BCaBA task list* (5th ed.). Retrieved June 22, 2017, from https://bacb.com/bcba-bcaba-task-list-5th-ed

Behavior Analyst Certification Board. (2014). *Professional and ethical compliance code for behavior analysts.* Retrieved June 22, 2017, from www.bacb.com/ethics-code/

Bejarano, R., & Hackenberg, T. D. (2007). IRT-stimulus contingencies in chained schedules: Implications for the concept of conditioned reinforcement. *Journal of the Experimental Analysis of Behavior, 88*, 215–227.

Belfiore, P., Basile, S., & Lee, D. (2008). Using a high probability command sequence to increase classroom compliance: The role of behavioral momentum. *Journal of Behavioral Education, 17*, 160–171. doi:10.1007/s10864-007-9054-x

Belfiore, P. J., Fritts, K. M., & Herman, B. C. (2008). The role of procedural integrity: Using self-monitoring to enhance discrete trial instruction. *Focus on Autism and Other Developmental Disabilities, 25*, 95–102.

Belisle, J., Dixon, M. R., Stanley, C. R., Munoz, B., & Daar, J. H. (2016). Teaching foundational perspective-taking skills to children with autism using the PEAK-T curriculum: Single-reversal "I–You" deictic frames. *Journal of Applied Behavior Analysis, 49*, 965–969.

Bell, M. C., & McDevitt, M. A. (2014). *Conditioned reinforcement.* In F. K. McSweeney & E. S. Murphy (Eds.), *Wiley-Blackwell handbook of operant and classical conditioning* (pp. 221–248). Oxford: Wiley-Blackwell.

Bellack, A. S., & Hersen, M. (1977). *Behavior modification: An introductory textbook.* New York: Oxford University Press.

Bellamy, G. T., Horner, R. H., & Inman, D. P. (1979). *Vocational habilitation of severely retarded adults.* Austin, TX: PRO-ED.

Bellini, S., & Akullian, J. (2007). A meta-analysis of video modeling and video self-modeling interventions for children and adolescents with autism spectrum disorders. *Exceptional Children, 73*(3), 264–287.

Bender, W. N., & Mathes, M. Y. (1995). Students with ADHD in the inclusive classroom: A hierarchical approach to strategy selection. *Intervention in School & Clinic, 30*(4), 226–234.

BenitezSantiago, A., & Miltenberger, R. G. (2016). Using video feedback to improve martial arts performance. *Behavioral Interventions, 31*, 12–27. doi:10.1002/bin.1424

Bennett, K., & Cavanaugh, R. A. (1998). Effects of immediate self-correction, delayed self-correction, and no correction on the acquisition and maintenance of multiplication facts by a fourth-grade student with learning disabilities. *Journal of Applied Behavior Analysis, 31*, 303–306.

Bennett, M. K., Manal, H., & Van Houten, R. (2014). A comparison of gateway in-street sign configuration to other driver prompts to increase yielding to pedestrians at crosswalks. *Journal of Applied Behavior Analysis, 47*, 3–15.

Berens, N. M., & Hayes, S. C. (2007). Arbitrarily applicable comparative relations: Experimental evidence for a relational operant. *Journal of Applied Behavior Analysis, 40*, 45–71.

Bergstrom, C. T., & West, J. (2017a). *Calling bullshit: Data reasoning in a digital world. [Syllabus with links to readings].* Retrieved December 4, 2017, at http://callingbullshit.org/syllabus.html

Bergstrom, C. T., & West, J. (2017b). *Visualization: Misleading axes on graphs.* Retrieved March 26, 2017, at http://callingbullshit.org/tools/tools_misleading_axes.html

Best, L. A., Smith, L. D., & Stubbs, D. A. (2001). Graph use in psychology and other sciences. *Behavioural Processes, 54*, 155–165.

Betz, A. M., Higbee, T. S., & Pollard, J. S. (2010). Promoting generalization of mands for information used by young children with autism. *Research in Autism Spectrum Disorder, 4*, 501–508.

Bicard, D. F. & Neef, N. A. (2002). Effects of strategic versus tactical instructions on adaptation to changing contingencies in children with ADHD. *Journal of Applied Behavior Analysis, 35*, 375–389.

Bicard, D. F., Ervin, A., Bicard, S. C., & Baylot-Casey, L. (2012). Differential effects of seating arrangements on disruptive behavior of fifth grade students during independent seatwork. *Journal of Applied Behavior Analysis, 45*, 407–411.

Bicard, D. F., Lott, V., Mills, J., Bicard, S., & Baylot-Casey, L. (2012). Effects of text messaged self-monitoring on class attendance and punctuality of at-risk college student athletes. *Journal of Applied Behavior Analysis, 45*, 205–210. doi:10.1901/jaba.2012.45-205

Bickel, W. K., Odum, A. L., & Madden, G. J. (1999). Impulsivity and cigarette smoking: Delay discounting in current, never, and ex-smokers. *Psychopharmacology, 146*, 447–454.

Biglan, A. (2015). *The nurture effect: How the science of human behavior can improve our lives & our world.* Oakland, CA: New Harbinger Publications.

Biglan, A., & Glenn, S. S. (2013). Toward prosocial behavior and environments: Behavioral and cultural contingence in a public health framework. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. 225–275). Washington, DC: American Psychological Association.

Biglan, A., Layton, G. L., Jones, L. B., Hankins, M., & Rusby, J. C. (2013). The value of workshops on psychological flexibility for early childhood special education staff. *Topics in Early Childhood Special Education, 32*, 196–210.

Bijou, S. W. (1955). A systematic approach to an experimental analysis of young children. *Child Development, 26*, 161–168.

Bijou, S. W. (1957). Patterns of reinforcement and resistance to extinction in young children. *Child Development, 28*, 47–54.

Bijou, S. W. (1958). Operant extinction after fixed-interval schedules with young children. *Journal of the Experimental Analysis of Behavior, 1*, 25–29.

Bijou, S. W., & Baer, D. M. (1961). *Child development: Volume I: A systematic and empirical theory.* Norwalk, CT: Appleton-Century-Crofts.

Bijou, S. W., & Baer, D. M. (1965). *Child development II: Universal stage of infancy.* Englewood Cliffs, NJ: Prentice Hall.

Bijou, S. W., & Baer, D. M. (1965). *Child development: Vol. 2. Universal stage of infancy.* New York: Appleton-Century-Crofts.

Bijou, S. W., & Baer, D. M. (1967). *Child development: Reading in experimental analysis.* New York: Meredith Publishing Company.

Bijou, S. W., Birnbrauer, J. S., Kidder, J. D., & Tague, C. (1966). Programmed instruction as an approach to teaching of reading, writing, and arithmetic to retarded children. *The Psychological Record, 16*, 505–522.

Bijou, S. W., Peterson, R. F., & Ault, M. H. (1968). A method to integrate descriptive and experimental field studies at the level of data and empirical concepts. *Journal of Applied Behavior Analysis, 1*, 175–191.

Billings, D. C., & Wasik, B. H. (1985). Self-instructional training with preschoolers: An attempt to replicate. *Journal of Applied Behavior Analysis, 18*, 61–67.

Billingsley, F., White, D. R., & Munson, R. (1980). Procedural reliability: A rationale and an example. *Behavioral Assessment, 2*, 247–256.

Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *The Behavior Analyst, 19*, 163–197.

Binder, L. M., Dixon, M. R., & Ghezzi, P. M. (2000). A procedure to teach self-control to children with attention-deficit/hyperactivity disorder. *Journal of Applied Behavior Analysis*, *33*, 233–237. doi:10.1901/jaba.2000.33-233

Birnbrauer, J. S. (1979). Applied behavior analysis, service, and the acquisition of knowledge. *The Behavior Analyst*, *2*, 15–21.

Birnbrauer, J. S. (1981). External validity and experimental investigation of individual behavior. *Analysis and Intervention in Developmental Disabilities*, *1*, 117–132.

Bishop, B. R., & Stumphauzer, J. S. (1973). Behavior therapy of thumb sucking in children: A punishment (time out) and generalization effect—what's a mother to do? *Psychological Reports*, *33*, 939–944.

Bishop, M. R., & Kenzer, A. L. (2012). Teaching behavioral therapists to conduct brief preference assessments during therapy sessions. *Research in Autism Spectrum Disorders*, *6*, 450–457.

Bjork, D. W. (1997). *B. F. Skinner: A life*. Washington, DC: American Psychological Association.

Blampied, N. M. (2013). Single-case research designs and the scientist-practitioner ideal in applied psychology. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. 177–197). Washington, DC: American Psychological Association.

Bloom, L. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.

Bloom, S. E., & Lambert, J. M. (2015). Implications for practice: Resurgence and differential reinforcement of alternative responding. *Journal of Applied Behavior Analysis*, *48*, 781–784. doi:10.1002/jaba.266

Bloom, S. E., Iwata, B. A., Fritz, J. N., Roscoe, E. M., & Carreau, A. B. (2011). Classroom application of a trial-based functional analysis. *Journal of Applied Behavior Analysis*, *44*, 19–31. doi:10.1901/jaba.2011.44-19

Boerke, K. W., & Reitman, D. (2011). Token economies. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 370–382). New York: Guilford Press.

Bolin, E. P., & Goldberg, G. M. (1979). Behavioral psychology and the Bible: General and specific considerations. *Journal of Psychology and Theology*, *7*, 167–175.

Bondy, A. (2012). The unusual suspects: Myths and misconceptions associated with PECS. *The Psychological Record*, *62*, 789–816.

Bondy, A., & Frost, L. (2002). *The Picture Exchange Communication System*. Newark, DE: Pyramid Educational Products.

Bondy, A. H., & Tincani, M. (2018). Effects of response cards on students with autism spectrum disorder or intellectual disability. *Education and Training in Autism and Developmental Disabilities*, *53*(1), 59–72.

Bones, R., Keenan, M., & McColgan, E. (2016). Men have breasts too: A pilot study using behavioural technology for skills-based training in detecting lumps in breasts. *Specialty Pedagogy*, *5*, 32–46.

Bord, S., Sidener, T. M., Reeve, K. F., & Disener, D. W. (2017). Teaching on-task rollerblading and ice-skating to a child with autism. *Behavior Analysis in Practice*, *10*(2), 178–182. doi:10.1007/s40617-016-0150-z

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2009). *Introduction to meta-analysis*. Hoboken, NJ: Wiley.

Bornstein, M. H., & Arterberry, M. E. (2010). The development of object categorization in young children: Hierarchical inclusiveness, age, perceptual attribute, and group versus individual analyses. *Developmental Psychology*, *46*, 350–365.

Bornstein, P. H., & Quevillon, R. P. (1976). The effects of a self-instructional package on overactive preschool boys. *Journal of Applied Behavior Analysis*, *9*, 179–188.

Borrero, C. S., Vollmer, T. R., Borrero, J. C., Bourret, J. C., Sloman, K. N., Samaha, A. L., & Dallery, J. (2010). Concurrent reinforcement schedules for problem behavior and appropriate behavior: Experimental applications of the matching law. *Journal of Experimental Analysis of Behavior*, *93*(3), 455–469. doi.org/10.1901/jeab.2010.93-455

Borrero, C. S. W., & Borrero, J. C. (2008). Descriptive and experimental analyses of potential precursors to problem behavior. *Journal of Applied Behavior Analysis*, *41*, 83–96. doi:10.1901/jaba.2008.41-83

Borrero, J. C., Vollmer, T. R., Wright, C. S., Lerman, D. C., & Kelley, M. E. (2002). Further evaluation of the role of protective equipment in the functional analysis of self-injurious behavior. *Journal of Applied Behavior Analysis*, *35*, 69–72.

Bosch, S., & Fuqua, R.W. (2001). Behavioral cusps: A model for selecting target behaviors. *Journal of Applied Behavior Analysis*, *34*, 123–125.

Bosch, S., & Hixson, M. D. (2004). The final piece to a complete science of behavior: Behavior development and behavioral cusps. *The Behavior Analyst Today*, *5*(3), 244–254.

Bouck, E. C., Satsangi, R., Muhl, A., & Bartlett, W. (2013). Using audio recorders to promote independence in grocery shopping for students with intellectual disability. *Journal of Special Education Technology*, *28*(4), 15–26.

Bourgeois, M. S. (1993). Effects of memory aids on the dyadic conversations of individuals with dementia. *Journal of Applied Behavior Analysis*, *26*, 77–87. doi:10.1901/jaba.1993.26-77

Bourret, J., Vollmer, T. R., & Rapp, J. T. (2004). Evaluation of a vocal mand assessment and vocal mand procedures. *Journal of Applied Behavior Analysis*, *37*, 129–144.

Bouton, M. E., & Nelson, J. E. (1998). The role of context in classical conditioning: Some implications for cognitive behavior therapy. In W. O'Donohue (Ed.), *Learning and behavior therapy* (pp. 59–84). Boston: Allyn and Bacon.

Boutot, E. A., & Digangi , S. A. (2018). Effects of activation of preferred stimulus on tummy time behavior of an infant with Down syndrome and associated hypotonia. *Behavior Analysis in Practice*, *11*(2), 144–147. https://doi.org/10.1007/s40617-018-0212-5

Bovi, G. M. D., Vladescu, J. C., DeBar, R. M., Carroll, R. A., & Sarokoff, R. A. (2017). Using video modeling with voice-over instruction to train public school staff to implement a preference assessment. *Behavior Analysis in Practice*, *10*(1), 72–76. doi 10.1007/s40617-016-0135-y

Bowman, L. G., Hardesty, S. L., & Mendres-Smith, A. E. (2013). A functional analysis of crying. *Journal of Applied Behavior Analysis*, *46*, 317–321. doi:10.1002/jaba.4

Bowman, L. G., Piazza, C. C., Fisher, W., Hagopian, L. P., & Kogan, J. S. (1997). Assessment of preference for varied versus constant reinforcement. *Journal of Applied Behavior Analysis*, *30*, 451–458.

Bowman-Perrott, L. D., Burke, M. D., deMarin, S., Zhang, N., & Davis, H. (2015). A meta-analysis of single-case research on behavior contracts: Effects on behavioral and academic outcomes among children and youth. *Behavior Modification*, *39*(2), 247–269.

Boyle, M. E., & Greer, R. D. (1983). Operant procedures and the comatose patient. *Journal of Applied Behavior Analysis*, *16*, 3–12.

Braam, S. J., & Poling, A. (1983). Development of intraverbal behavior in mentally retarded individuals through transfer of stimulus control procedures: Classification of verbal responses. *Applied Research in Mental Retardation*, *4*, 279–302.

Braam, S. J., & Sundberg, M. L. (1991). The effects of specific versus nonspecific reinforcement on verbal behavior. *The Analysis of Verbal Behavior*, *9*, 19–28.

Bradley, K. P., & Poling, A. (2010). Defining delayed consequences as reinforcers: Some do, some don't, and nothing changes. *The Analysis of Verbal Behavior*, *26*, 41–49.

Brady, K. K., & Kubina, Jr., R. M. (2010). Endurance of multiplication fact fluency for students with attention deficit hyperactivity disorder. *Behavior Modification*, *34*(2), 79–93.

Braem, M. D., & Mills, D. S. (2010). Factors affecting response of dogs to obedience instruction: A field and experimental study. *Applied Animal Behaviour Science*, *125*, 47–55.

Brame, P., Heward, W. L., & Bicard, S. C. (2002). *Effects of a story facts recall game on students' off-task behavior during sustained silent reading (SSR) and retention of story facts*. Paper presented at the 28th Annual Convention of the Association for Behavior Analysis International, Toronto, Canada.

Brame, P. B. (2001). *Making sustained silent reading (SSR) more effective: Effects of a story fact recall game on students' off-task behavior during SSR and retention of story facts*. Unpublished doctoral dissertation, The Ohio State University, Columbus, OH.

Branch, M. N. (2014). Malignant side effects of null-hypothesis significance testing. *Theory & Psychology*, *24*(2), 256-277. doi:10.1177/0959354314525282

Branch, M. N., & Pennypacker, H. S. (2013). Generality and generalization of research findings. In G. J. Madden, W. V. Dube, T. D. Hackenberg, G. P. Hanley, & K. A. Lattal (Eds.), *APA Handbook of Behavior Analysis, Vol. 1: Methods and Principles* (pp. 151–175). Washington, DC: American Psychological Association.

Brantley, D. C., & Webster, R. E. (1993). Use of an independent group contingency management system in a regular classroom setting. *Psychology in the Schools*, *30*, 60–66.

Brantner, J. P., & Doherty, M. A. (1983). A review of timeout: A conceptual and methodological analysis. In S. Axelrod & J. Apsche (Eds.), *The effects of punishment on human behavior* (pp. 87–132). New York: Academic Press.

Brassil, N., Hyland, J., O'Hora, D., & Stewart, I. (2018). Reversing time and size: Mutual entailment of nonarbitrary temporal and magnitude relational responding. *The Psychological Record*. https://doi.org/10.1007/s40732-018-0323-y

Brazeau, K., Rehfeldt, R. A., Mazo, A., Smalley, S., Krus, S., & Henson, L. (2017). On the efficacy of mindfulness, defusion, and behavioral skills training on job interviewing skills in dually-diagnosed adults with developmental disorders. *Journal of Contextual Behavior Science*, *6*, 145–151.

Brazelton, T. B., & Sparrow, J. D. (2006). *Touchpoints*. Cambridge, MA: Da Capo Books.

Brethower, D. C., & Reynolds, G. S. (1962). A facilitative effect of punishment on unpunished behavior. *Journal of the Experimental Analysis of Behavior*, *5*, 191–199.

Brigham, T. A. (1980). Self-control revisited: Or why doesn't anyone read Skinner anymore? *The Behavior Analyst*, *3*, 25–33.

Brigham, T. A. (1983). Self-management: A radical behavioral perspective. In P. Karoly & F. H. Kanfer (Eds.), *Self-management and behavior change: From theory to practice* (pp. 32–59). New York: Pergamon Press.

Brinton, L. J., & Brinton, D. M. (2010). *The linguistic structure of Modern English* (2nd ed.). Amsterdam, Netherlands: John Benjamins Publishing Company.

Brobst, B., & Ward, P. (2002). Effects of public posting, goal setting, and oral feedback on the skills of female soccer players. *Journal of Applied Behavior Analysis*, *27*, 247–257.

Broden, M., Hall, R. V., & Mitts, B. (1971). The effect of self-recording on the classroom behavior of two eighth-grade students. *Journal of Applied Behavior Analysis*, *4*, 191–199.

Brodhead, M. (2015). Maintaining professional relationships in an interdisciplinary setting: Strategies for navigating non-behavioral treatment recommendations for individuals with autism. *Behavior Analysis in Practice*, *8*, 70–78. Published online, 10 Feb 2015. doi 10.1007/s40617-015-0042-7

Brodhead, M. T., & Higbee, T. S. (2012). Teaching and maintaining ethical behavior in a professional organization. *Behavior Analysis in Practice*, *5*, 82–88.

Brodhead, M. T., Al-Dubayan, M. N., Mates, M., Abel, E. A., & Brouwers, L. (2016). An evaluation of a brief video-based multiple-stimulus without replacement preference assessment. *Behavior Analysis in Practice*, *9*(2), 160–164. doi:10.1007/s40617-015-0081-0

Brodhead, M. T., Durán, L., & Bloom, S. E. (2014). Cultural and linguistic diversity in recent verbal behavior research on individuals with disabilities: A review and implications for research and practice. *The Analysis of Verbal Behavior*, *30*, 75– 86. doi.org/10.1007/s40616-014-0009-8

Brookhart, S. M., & Nitko. A. J. (2019). *Educational assessment of students* (8th ed.). Upper Saddle River, NJ: Pearson.

Brosnan, J., & Healy, O. (2011). A review of behavioral interventions for the treatment of aggression in individuals with developmental disabilities. *Research in Developmental Disabilities*, *32*, 437–446. doi:10.1016/j.ridd.2010.12.023

Brossart, D. F., Vannest, K. J., Davis, J. L., & Patience, M. A. (2014). Incorporating nonoverlap indices with visual analysis for quantifying intervention effectiveness in single-case experimental designs. *Neuropsychological Rehabilitation*, *24*(3-4), 464–491. doi.org/10.1080/09602011.2013.868361

Brothers, K. J., Krantz, P. J., & McClannahan, L. E. (1994). Office paper recycling: A function of container proximity. *Journal of Applied Behavior Analysis*, *27*, 153–160.

Browder, D. M., Ahlgrim-Delzell, L., Spooner, F., Mims, P. J., & Baker, J. N. (2009). Using time delay to teach literacy to students with severe developmental disabilities. *Exceptional Children*, *75*, 343–364.

Brown, F. E., McDonnell, J. J., & Snell, M. E. (2016). *Instruction of students with severe disabilities* (8th ed.). Upper Saddle River, NJ: Pearson.

Brown, K. A., Wacker, D. P., Derby, K. M., Peck, S. M., Richman, D. M., Sasso, . . . Harding, J. W. (2000). Evaluating the effects of functional communication training in the presence and absence of establishing operations. *Journal of Applied Behavior Analysis*, *33*, 53–71.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Brown, S. A., Dunne, J. D., & Cooper, J. O. (1996) Immediate retelling's effect on student retention. *Education and Treatment of Children*, *19*, 387–407.

Browning, R. M. (1967). A same-subject design for simultaneous comparison of three reinforcement contingencies. *Behavior Research and Therapy*, *5*, 237–243.

Bruhn, A., McDonald, S., & Kreigh, C. (2015). Self-monitoring interventions for students with behavior problems: A systematic review of current research. *Behavioral Disorders*, *40*, 102–121.

Buchanan, J. A., & Fisher, J. E. (2002). Functional assessment and noncontingent reinforcement in the treatment of disruptive vocalization in elderly dementia patients. *Journal of Applied Behavior Analysis*, *35*, 99–103.

Budd, K. S., & Baer, D. M. (1976). Behavior modification and the law: Implications of recent judicial decisions. *Journal of Psychiatry and Law*, *4*, 171–244.

Budd, K. S., & Stokes, T. F. (Eds.). (2003). *A small matter of proof: The legacy of Donald M. Baer*. Reno, NV: Context Press.

Buell, J. S., Stoddard, P., Harris, F. R., & Baer, D. M. (1968). Collateral social development accompanying reinforcement of outdoor play in a preschool child. *Journal of Applied Behavior Analysis*, *1*, 167–173.

Bulla, A. J., & Frieder, J. E. (2017). Self-and-Match system suppresses vocal stereotypy during independent work. *Behavior Analysis Research and Practice*, *17*, 274–285.

Bullock, C. E., & Hackenberg, T. D. (2006). Second-order schedules of token reinforcement with pigeons: Implications for unit price. *Journal of the Experimental Analysis of Behavior*, *85*, 95–106.

Bullock, C. E., Fisher, W. W., & Hagopian, L. P. (2017). Description and validation of a computerized behavioral data program: "BDataPro." *The Behavior Analyst*, *40*, 275–285.

Burch, M. R., & Bailey, J. S. (1999). *How dogs learn*. New York: Howell.

Burckhardt, R., Manicavasagar, V., Batterham, P. J., Hadzi-Pavlovic, D., & Shand, F. (2017). Acceptance and commitment therapy universal prevention program for adolescents: A feasibility study. *Child and Adolescent Psychiatry and Mental Health*, *11*, 27.

Burger, K. S., Fisher, J. O., & Johnson, S. L. (2011). Mechanisms behind the portion size effect: Visibility and bite size. *Obesity*, *19*, 546–551.

Burgio, L. D., Whitman, T. L., & Johnson, M. R. (1980). A self-instructional package for increasing attending behavior in educable mentally retarded children. *Journal of Applied Behavior Analysis*, *13*, 443–459.

Busacca, M. L., Moore, D. W., & Anderson, A. (2015). Self-management for primary school students demonstrating problem behavior in regular classrooms: Evidence review of single-case design research. *Journal of Behavioral Education*, *24*, 373–401.

Bush, K. M. (1993). Stimulus equivalence and cross-modal transfer. *The Psychological Record*, *43*, 567–584.

Bush, K. M., Sidman, M., & de Rose, T. (1989). Contextual control of emergent equivalence relations. *Journal of the Experimental Analysis of Behavior*, *51*, 29–45.

Bushell, D., Jr., & Baer, D. M. (1994). Measurably superior instruction means close, continual contact with the relevant outcome data. Revolutionary! In R. Gardner III, D. M. Sainato, J. O. Cooper, T. E. Heron, W. L. Heward, J. Eshleman, & T. A. Grossi (Eds.), *Behavior analysis in education: Focus on measurably superior instruction* (pp. 3–10). Pacific Grove, CA: Brooks/Cole.

Butryn, M. L., Forman, E., Hoffman, K., Shaw, J., & Juarascio, A. (2011). A pilot study of acceptance and commitment therapy for promotion of physical activity. *Journal of Physical Activity and Health*, *8*, 516–522.

Byrd, M. R., Richards, D. F., Hove, G., & Friman, P. C. (2002). Treatment of early onset hair pulling as a simple habit. *Behavior Modification*, *26*(3), 400–411.

Byrne, B. L., Rehfeldt, R. A., & Aguirre, A. A. (2014). Evaluating the effectiveness of the stimulus pairing observation procedure and multiple exemplar instruction on tact and listener responses in children with autism. *The Analysis of Verbal Behavior*, *30*, 160–169. doi.org/10.1007/s40616- 014-0020-0

Byrne, T., LeSage, M. G., & Poling, A. (1997). Effects of chlorpromazine on rats' acquisition of lever-press responding with immediate and delayed reinforcement. *Pharmacology Biochemistry and Behavior*, *58*, 31–35.

Bzoch, K. R., League, R., & Brown, V. L. (2003). Receptive-expressive emergent language test (3rd ed.). Austin, TX: PRO-ED.

Cachia, R. L., Anderson, A., & Moore, D. W. (2016). Mindfulness, stress and well-being in parents of children with autism spectrum disorder: A systematic review. *Journal of Child and Family Studies*, *25*, 1–14.

Cakiroglu, O. (2014). Effects of preprinted response cards on rates of academic response, opportunities to respond, and correct on academic responses of students with mild intellectual disability. *Journal of Intellectual & Developmental Disability, 39*(1), 73–85.

Caldwell, N. K., Wolery, M., Werts, M. G., & Caldwell, Y. (1996). Embedding instructive feedback into teacher-student interactions during independent seatwork. *Journal of Behavioral Education, 6*, 459–480.

Cammilleri, A. P., & Hanley, G. P. (2005). Use of a lag differential reinforcement contingency to increase varied selections of classroom activities. *Journal of Applied Behavior Analysis, 38*, 111–115.

Cammilleri, A. P., Tiger, J. H., & Hanley, G. P. (2008). Developing stimulus control of young children's requests to teachers: Classwide applications of multiple schedules. *Journal of Applied Behavior Analysis, 41*, 299–303. doi:10.1901/jaba.2008.41-299

Campbell, A., & Anderson, C. M. (2011). Check-in/Check-out: A systematic evaluation and component analysis. *Journal of Applied Behavior Analysis, 44*, 315–326.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* New York: Houghton Mifflin Company.

Campbell, R. C., & Stremel-Campbell, K. (1982). Programming "loose training" as a strategy to facilitate language generalization. *Journal of Applied Behavior Analysis, 15*, 295–301.

Cannella-Malone, H. I., Brooks, D. G., & Tullis, C. A. (2013). Using self-directed video prompting to teach students with intellectual disabilities. *Journal of Behavioral Education, 22*, 169–189.

Cannella-Malone, H. I., Chan, J. M., & Jimenez, E. D. (2017). Comparing self-directed video prompting to least-to-most prompting in post-secondary students with moderate intellectual disabilities. *International Journal of Developmental Disabilities, 63*, 211–220. doi.org/10.1080/20473869.2017.1301695

Cannella-Malone, H. I., Fleming, C., Chung, Y. C., Wheeler, G. M., Basbagill, A. R., & Singh, A. H. (2011). Teaching daily living skills to seven individuals with severe intellectual disabilities: A comparison of video prompting to video modeling. *Journal of Positive Behavior Interventions, 13*(3), 144–153. doi.org/10.1177/1098300710366593

Capriotti, M. R., Turkel, J. E., Johnson, R. A., Espil, F. M., & Woods, D. W. (2017). Comparing fixed-amount and progressive-amount DRO schedules for tic suppression in youth with chronic tic disorders. *Journal of Applied Behavior Analysis, 50*, 106–120.

Carbone, V. J. (2013). The establishing operation and teaching verbal behavior. *The Analysis of Verbal Behavior, 29*, 45–49.

Carbone V. J., Morgenstern, B., Zecchin-Tirri, G., & Kolberg, L. (2010). The role of the reflexive conditioned motivating operation (CMO-R) during discrete trial instruction of children with autism. *Focus on Autism and Other Developmental Disabilities, 25*, 110–124.

Carbone, V. J., Sweeney-Kervin, E., Attanasio, V., & Kasper, T. (2010). Increasing the vocal responses of children with autism and developmental disabilities using manual sign mand training and prompt delay. *Journal of Applied Behavior Analysis, 43*(4), 705–709.

Cariveau, T., & Kodak, T. (2017). Programming randomized dependent group contingency and common stimuli to promote durable behavior change. *Journal of Applied Behavior Analysis, 50*, 121–133. doi:10.1002/jaba.352

Cariveau, T., Kodak, T., & Campbell, V. (2016). The effects of intertrial interval and instructional format on skill acquisition and maintenance for children with autism spectrum disorders. *Journal of Applied Behavior Analysis, 49*, 809–825. doi:10.1002/jaba.322

Carlson, J. F., Geisinger, K. F., & Jonson, J. L. (Eds.). (2017). *Twentieth mental measurements yearbook.* Lincoln, NE: Buros Center for Testing.

Carp, C. L., & Petursdottir, A. I. (2015). Intraverbal naming and equivalence class formation in children. *Journal of the Experimental Analysis of Behavior, 104*, 223–240. doi:10.1002/jeab.183

Carr, D. (2003). Effects of exemplar training in exclusion responding of auditory-visual discrimination tasks with children with autism. *Journal of Applied Behavior Analysis, 36*, 507–524.

Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis, 18*, 111–126.

Carr, E. G., & Kologinsky, E. (1983). Acquisition of sign language by autistic children II: Spontaneity and generalization effects. *Journal of Applied Behavior Analysis, 16*, 297–314.

Carr, E. G., Newsom, C. D., & Binkoff, J. A. (1976). Stimulus control of self-destructive behavior in a psychotic child. *Journal of Abnormal Child Psychology, 4*, 139–153.

Carr, J. E. (2005). Recommendations for reporting multiple-baseline designs across participants. *Behavioral Interventions, 20*, 219–224.

Carr, J. E., & Miguel, C. F. (2012). The analysis of verbal behavior and its therapeutic applications. In G. J. Madden (Ed.), *APA Handbook of Applied Behavior Analysis* (pp. 329–352). *Vol. 2*. Washington, DC: American Psychological Association.

Carr, J. E., Kellum, K. K., & Chong, I. M. (2001). The reductive effects of noncontingent reinforcement: Fixed-time versus variable-time schedules. *Journal of Applied Behavior Analysis, 34*, 505–509.

Carr, J. E., Nicolson, A. C., & Higbee, T. S. (2000). Evaluation of a brief multiple-stimulus preference assessment in a naturalistic context. *Journal of Applied Behavior Analysis, 33*, 353–357.

Carr, J. E., Ratcliff, C. L., Nosik, M. R., & Johnston, J. M. (in press). Professional certification for practitioners of applied behavior analysis. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (2nd ed.). New York: Guilford Press.

Carr, M. E., Moore, D. W., & Anderson, A. (2014). Self-management interventions on students with autism: A meta-analysis of single-subject research. *Exceptional Children, 81*, 28–44.

Carr, S. C., & Punzo, R. P. (1993). The effects of self-monitoring of academic accuracy and productivity on the performance of students with behavioral disorders. *Behavioral Disorders, 18*, 241–250.

Carrigan, P. F., Jr., & Sidman, M. (1992). Conditional discrimination and equivalence relations: A theoretical analysis of control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 58*, 459–504.

Carroll, R. A., & St. Peter, C. C. (2014). Methods for assessing social validity of behavioral intervention plans for children with attention deficit hyperactivity disorder. *Acta de Investigación Psicológica, 4*(3), 1642–1656.

Carroll, R. A., Joachim, B. T., St. Peter, C. C., & Robinson, N. (2015). A comparison of error-correction procedures on skill acquisition during discrete-trial instruction. *Journal of Applied Behavior Analysis, 48*, 257–273.

Carroll, R. A., Kodak, T., & Fisher, W. W. (2013). An evaluation of programmed treatment integrity errors during discrete trial instruction. *Journal of Applied Behavior Analysis, 46*, 379–394.

Carroll, R. A., Rapp, J. T., Colby-Dirksen, A. M., & Lindenberg, A. M. (2009). Detecting changes in simulated events II: Using variations of momentary time sampling to measure changes in duration events. *Behavioral Interventions, 24*, 137–155.

Carter, D. E., & Eckerman, D. A. (1975). Symbolic matching by pigeons: Rate of learning complex discriminations predicted from simple discriminations. *Science, 187*, 662–664.

Carter, M., & Grunsell, J. (2001). The behavior chain interruption strategy: A review of research and discussion of future directions. *Journal of The Association for Persons with Severe Handicaps, 26*(1), 37–49.

Carter, S. L. (2007). Review of recent treatment acceptability research. *Education and Training in Developmental Disabilities, 42*(3), 301–316.

Cassella, M. D., Sidener, T. M., Sidener, D. W., & Progar, P. R. (2011). Response interruption and redirection for vocal stereotypy in children with autism: A systematic replication. *Journal of Applied Behavior Analysis, 44*, 169–173. doi:10.1901/jaba.2011.44-169

Cassidy, S., Roche, B., Colbert, D., Stewart, I., & Grey, I. M. (2016). A relational frame skills training intervention to increase general intelligence and scholastic aptitude. *Learning and Individual Differences, 47*, 222–235.

Cassidy, S., Roche, B., & O'Hora, D. (2010). Relational Frame Theory and human intelligence. *European Journal of Behavior Analysis, 11*, 37–51.

Castro, M., Rehfeldt, R. A., & Root, W. B. (2016). On the role of values clarification and committed actions in enhancing the engagement of direct care workers with clients with severe developmental disorders. *Journal of Contextual Behavioral Science, 5*, 201–207.

Cataldo, M. F. (2017). Driving with the rear view mirror. *The Behavior Analyst, 40*, 161–165. doi:10.1007/s40614-017-0103-z

Catania, A. C. (1973). The concept of the operant in the analysis of behavior. *Behaviorism, 1*, 103–116.

Catania, A. C. (1975). The myth of self-reinforcement. *Behaviorism, 3*, 192–199.

Catania, A. C. (1979). *Learning.* Upper Saddle River, NJ: Prentice Hall.

Catania, A. C. (1993). What John B. Watson left out of his behaviorism. *Mexican Journal of Behavior Analysis, 19*, 133–146.

Catania, A. C. (1998). *Learning* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.

Catania, A. C. (2011). Basic operant contingencies: Main effects and side effects. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 34–54). New York: Guilford Press.

Catania, A. C. (2013). *Learning* (5th ed.) Cornwall-on-Hudson, NY: Sloan Publishing.

Catania, A. C., & Harnad, S. (Eds.). (1988). *The selection of behavior: The operant behaviorism of B. F. Skinner: Comments and controversies*. New York: Cambridge University Press.

Causin, K. G., Albert, K. M., Carbone V. J., & Sweeney-Kerwin, E. J. (2011). The role of joint control in teaching listener responding to children with autism and other developmental disabilities. *Research in Autism Spectrum Disorders, 7*, 997–1011.

Cautela, J. R. (1971). Covert conditioning. In A. Jacobs & L. B. Sachs (Eds.), *The psychology of private events: Perspective on covert response systems* (pp. 109–130). New York: Academic Press.

Cavalier, A., Ferretti, R., & Hodges, A. (1997). Self-management within a classroom token economy for students with learning disabilities. *Research in Developmental Disabilities, 18*(3), 167–178.

Cavanaugh, R. A., Heward, W. L., & Donelson, F. (1996). Effects of response cards during lesson closure on the academic performance of secondary students in an earth science course. *Journal of Applied Behavior Analysis, 29*, 403–406.

Celiberti, D. A., Bobo, H. E., Kelly, K. S., Harris, S. L., & Handleman, J. S. (1997). The differential and temporal effects of antecedent exercise on the self-stimulatory behavior of a child with autism. *Research in Developmental Disabilities, 18*, 139–150.

Cenghar, M., Budd, A., Farrell, N., & Fienup, D. M. (2018). A review of prompt-fading procedures: Implications for effective and efficient skills acquisition. *Journal of Developmental and Physical Disabilities, 30*, 155–173. doi.org/10.1007/s10882-017-9575-8

Chadsey-Rusch, J., Drasgow, E., Reinoehl, B., Halle, J., & Collet-Klingenberg, L. (1993). Using general-case instruction to teach spontaneous and generalized requests for assistance to learners with severe disabilities. *Journal of the Association for Persons with Severe Handicaps, 18*, 177–187.

Challies, D. M., Hunt, M., Garry, M., & Harper, D. N. (2011). Whatever gave you that idea? False memories following equivalence training: A behavioral account of the misinformation effect. *Journal of the Experimental Analysis of Behavior, 96*, 343–362.

Chance, P., & Heward, W. L. (2010). Climate change: Meeting the challenge. *The Behavior Analyst, 33*, 145–151.

Chancey, C., Weihl, C., Root, W. B., McCauley, D., Takeguchi, K., & Rehfeldt, R. A. (in press). The impact of mindfulness skills on interactions between direct care staff and adults with developmental disabilities. *Journal of Contextual Behavioral Science*.

Charlop, M. H., Burgio, L. D., Iwata, B. A., & Ivancic, M. T. (1988). Stimulus variation as a means of enhancing punishment effects. *Journal of Applied Behavior Analysis, 21*, 89–95.

Charlop-Christy, M. H., & Carpenter, M. H. (2000). Modified incidental teaching sessions: A procedure for parents to increase spontaneous speech in their children with autism. *Journal of Positive Behavioral Interventions, 2*, 98–112.

Charlop-Christy, M. H., & Haymes, L. K. (1998). Using objects of obsession as token reinforcers for children with autism. *Journal of Autism and Developmental Disorders, 28*(3), 189–198.

Charlop-Christy, M. H., Carpenter, M., Le, L., LeBlanc, L. A., & Kellet, K. (2002). Using the picture exchange communication system (PECS) with children with autism: Assessment of PECS acquisition, speech, social-communicative behavior, and problem behavior. *Journal of Applied Behavioral Analysis, 35*, 213–231.

Charlop-Christy, M. H., Le, L., & Freeman, K. A. (2000). A comparison of video modeling with in vivo modeling for teaching children with autism. *Journal of Autism and Developmental Disorders, 30*(6), 537–552.

Chase, P. N. (2006). Teaching the distinction between positive and negative reinforcement. *The Behavior Analyst, 29*, 113–115.

Chase, P. N., & Danforth, J. S. (1991). The role of rules in concept learning. In L. J. Hayes and P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 205–225). Reno, NV: Context Press.

Chawla, N., & Ostafin, B. (2007). Experiential avoidance as a functional dimensional approach to psychopathology: An empirical review. *Journal of Clinical Psychology, 63*, 871–890.

Chiang, S. J., Iwata, B. A., & Dorsey, M. F. (1979). Elimination of disruptive bus riding behavior via token reinforcement on a "distance-based" schedule. *Education and Treatment of Children, 2*, 101–109.

Chiesa, M. (1994). *Radical behaviorism: The philosophy and the science*. Boston: Authors Cooperative.

Christensen, A. M., Lignugaris-Kraft, B., & Fiechtl, B. (1996). Teaching pairs of preschoolers with disabilities to seek adult assistance in response to simulated injuries: Acquisition and promotion of observational learning. *Education & Treatment of Children, 19*, 3–18.

Christian, L., & Poling, A. (1997). Using self-management procedures to improve the productivity of adults with developmental disabilities in a competitive employment setting. *Journal of Applied Behavior Analysis, 30*, 169–172.

Christle, C. A., & Schuster, J. W. (2003). The effects of using response cards on student participation, academic achievement, and on-task behavior during whole-class, math instruction. *Journal of Behavioral Education, 12*, 147–165.

Ciccone, F. J., Graff, R. B., & Ahearn, B. H. (2005). An alternate scoring method for the multiple stimulus without replacement preference assessment. *Behavioral Interventions, 20*, 121–127.

Cihak, D. F. (2009). Using video modeling via handheld computers to improve geometry skills for high school students with learning disabilities. *Journal of Special Education Technology, 26*(1), 1–19. doi.org/10.1177/016264340902400402

Cipani, E., Brendlinger, J., McDowell, L., & Usher, S. (1991). Continuous vs. intermittent punishment: A case study. *Journal of Developmental and Physical Disabilities, 3*, 147–156.

Cipani, E., Robinson, S., & Toro, H. (2003). *Ethical and risk management issues in the practice of ABA*. Paper presented at the annual conference of the Florida Association of Behavior Analysis, St. Petersburg, FL.

Clark, H. B., Rowbury, T., Baer, A., & Baer, D. M. (1973). Time out as a punishing stimulus in continuous and intermittent schedules. *Journal of Applied Behavior Analysis, 6*, 443–455.

Clayton, M. C., & Hayes, L. J. (2004). A comparison of match-to-sample and respondent-type training of equivalence classes. *The Psychological Record, 54*, 579–602.

Clemons, L. L., Mason, B. A., Garrison-Kane, L., & Wills, H. P. (2016). Self-monitoring for high-school students with disabilities: A cross-categorical investigation of I-Connect. *Journal of Positive Behavior Interventions, 18*, 145–155. doi:10.1177/1098300715596134

Cleveland, W. S. (1984b). Graphs in scientific publications. *The American Statistician, 38*, 261–269.

Clevenger, T. M., & Graff, R. B. (2005). Assessing object-to-picture and picture-to-object matching as prerequisite skills for pictorial preference assessments. *Journal of Applied Behavior Analysis, 38*, 543–547.

Codding, R. S., Feinberg, A. B., Dunn, E. K., & Pace, G. M. (2005). Effects of immediate performance feedback on implementation of behavior support plans. *Journal of Applied Behavior Analysis, 38*, 205–219.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. A. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37–46.

Cohen, M. J., & Gerhardt, P. F. (2015). *Visual supports for people with autism: A guide for parents and* professionals (2nd ed.). Baltimore: Brookes.

Cohen-Almeida, D., Graff, R. B., & Ahearn, W. H. (2000). A comparison of verbal and tangible stimulus preference assessments. *Journal of Applied Behavior Analysis, 33*, 329–334.

Colbert, D., Dobutowitsch, M., Roche, B., & Brophy, C. (2017). The proxy-measurement of intelligence quotients using a relational skills abilities index. *Learning & Individual Differences, 57*, 114–122.

Cole, D. M., & Witts, B. N. (2015). Formative graphing with a Microsoft Excel 2013 template. *Behavior Analysis: Research and Practice, 15*, 171–186.

Cole, G. A., Montgomery, R. W., Wilson, K. M., & Milan, M. A. (2000). Parametric analysis of overcorrection duration effects: Is longer really better than shorter? *Behavior Modification, 24*, 359–378.

Coleman-Martin, M. B., & Wolff Heller, K. (2004). Using a modified constant prompt-delay procedure to teach spelling to students with physical disabilities. *Journal of Applied Behavior Analysis, 37*, 469–480.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407–428.

Collins, F. L., Ricci, J. A., & Burkett, P. A. (1981). Behavioral training for myopia: Long term maintenance of improved acuity. *Behaviour Research and Therapy, 19*(3), 265–268.

Colquhoun, D. (2014). An investigation of the false discovery rate and the misinterpretation of *p*-values. *Royal Society Open Science, 1*: 140216. http://dx.doi.org/10.1098/rsos.140216

Colquhoun, D. (2016). The problem with *p*-values. *Aeon*. Retrieved June 15, 2018, from https://aeon.co/essays/it-s-time-for-science-to-abandon-the-term-statistically-significant?utm_source=Aeon+Newsletter&utm_campaign=fd9ca5ba57-Daily_Newsletter_11_October_201610_11_2016&utm_medium=email&utm_term=0_411a82e59d-fd9ca5ba57-68769545

Common, E. A., & Lane, K. L. (2017). Social validity assessment. In J. K. Luiselli (Ed.), *Applied behavior analysis advanced guidebook: A manual for professional practice* (pp. 73–89). London, UK: Academic Press.

Comunidad Los Horcones. (1992). Natural reinforcement to improve education. *Journal of Applied Behavior Analysis, 25*, 71–75.

Conaghan, B. P., Singh, N. N., Moe, T. L., Landrum, T. J., & Ellis, C. R. (1992). Acquisition and generalization of manual signs by hearing-impaired adults with mental retardation. *Journal of Behavioral Education, 2*, 175–203.

Connell, M. C., Carta, J. J., & Baer, D. M. (1993). Programming generalization of in-class transition skills: Teaching preschoolers with developmental delays to self-assess and recruit contingent teacher praise. *Journal of Applied Behavior Analysis, 26*, 345–352.

Conners, J., Iwata, B. A., Kahng, S. W., Hanley, G. P., Worsdell, A. S., & Thompson, R. H. (2000). Differential responding in the presence and absence of discriminative stimuli during multi-element functional analyses. *Journal of Applied Behavior Analysis, 33*, 299–308.

Constantine, B., & Sidman, M. (1975). The role of naming in delayed matching to sample. *American Journal of Mental Deficiency, 79*, 680–689.

Contreras, B. P., & Betz, A. M. (2016). Using lag schedules to strengthen the intraverbal repertoires of children with autism. *Journal of Applied Behavior Analysis, 49*, 3–16. doi:10.1002/jaba.271

Conyers, C., Miltenberger, R., Maki, A., Barenz, R., Jurgens, M., Sailer, A., . . . Koop, B. (2004). A comparison of response cost and differential reinforcement of other behaviors to reduce disruptive behavior in a preschool classroom. *Journal of Applied Behavior Analysis, 37*, 411–415.

Conyers, C., Miltenberger, R., Romaniuk, C., Kopp, B., & Himle, M. (2003). Evaluation of DRO schedules to reduce disruptive behavior in a preschool classroom. *Child and Family Behavior Therapy, 25*, 1–6.

Conyers, C., Miltenberger, R. G., Peterson, B., Gubin, A., Jurgens, M., Selders, A., . . . Barenz, R. (2004). An evaluation of in vivo desensitization and video modeling to increase compliance with dental procedures in persons with mental retardation. *Journal of Applied Behavior Analysis, 37*, 233–238.

Cook, J. L., Rapp, J. T., & Schulze, K. A. (2015). Differential negative reinforcement of other behavior to increase wearing of a medical bracelet. *Journal of Applied Behavior Analysis, 48*, 901–906. doi:10.1002/jaba.228

Cooke, N. L. (1984). Misrepresentations of the behavioral model in preservice teacher education textbooks. In W. L. Heward, T. E. Heron, D. S. Hill, & J. Trap-Porter (Eds.), *Focus on behavior analysis in education* (pp. 197–217). Columbus, OH: Charles E. Merrill Publishing Company.

Coon, J. T., & Miguel, C. F. (2012). The role of increased exposure to transfer-of-stimulus-control procedures on the acquisition of intraverbal behavior. *Journal of Applied Behavior Analysis, 45*, 657–666. doi:10.1901/jaba.2012.45-657

Cooper, H., & Patall, E. A. (2009). The relative benefits of meta-analysis conducted with individual participant data versus aggregated data. *Psychological Methods, 14*, 165–176.

Cooper, J. O. (1981). *Measuring behavior* (2nd ed.). Columbus, OH: Charles E. Merrill.

Cooper, J. O. (2005). Applied research: The separation of applied behavior analysis and precision teaching. In W. L. Heward, T. E. Heron, N. A. Neef, S. M. Peterson, D. M. Sainato, G. Cartledge, R. Gardner, III, L. D. Peterson, S. B. Hersh, & J. C. Dardig (Eds.), *Focus on behavior analysis in education: Achievements, challenges, and opportunities* (pp. 295–303). Upper Saddle River, NJ: Prentice Hall/Merrill.

Cooper, K. J., & Browder, D. M. (1997). The use of a personal trainer to enhance participation of older adults with severe disabilities in a community water exercise class. *Journal of Behavioral Education, 7*, 421–434.

Cope, J. G., Allred, L. J., & Morsell, J. M. (1991). Signs as deterrents of illegal parking in spaces designated for individuals. *Journal of Applied Behavior Analysis, 24*, 59–63.

Corey, G., Corey, M. S., & Callanan, P. (1993). *Issues and ethics in the helping professions* (4th ed.). Pacific Grove, CA: Brooks/Cole.

Correa, V., Poulson, C., & Salzburg, C. (1984). Training and generalization of reach-grasp behavior in blind, retarded young children. *Journal of Applied Behavior Analysis, 17*, 57–69.

Costenbader, V., & Reading-Brown, M. (1995). Isolation timeout used with students with emotional disturbance. *Exceptional Children, 61*(4), 353–364.

Council for Exceptional Children. (2009). *CEC's policy on physical restraint and seclusion procedures in school settings*. Arlington, VA: Author. Retrieved from https://www.cec.sped.org/~/media/Files/Policy/CEC%20Professional%20Policies%20and%20Positions/restraint%20and%20seclusion.pdf

Cowdery, G., Iwata, B. A., & Pace, G. M. (1990). Effects and side effects of DRO as treatment for self-injurious behavior. *Journal of Applied Behavior Analysis, 23*, 497–506.

Cowley, B. J., Green, G., & Braunling-McMorrow, D. (1992). Using stimulus equivalence procedures to teach name-face matching to adults with brain injuries. *Journal of Applied Behavior Analysis, 25*, 461–475.

Cox, A. L., Gast, D. L., Luscre, D., & Ayers, K. M. (2009). The effects of weighted vests on appropriate in-seat behaviors of elementary-age students with autism and severe to profound intellectual disabilities. *Focus on Autism and Other Developmental Disabilities, 24*, 17–26.

Coy, J. N., & Kostewicz, D. E. (2018). *Teaching Exceptional Children, 50*, 301–309. doi:10.1177/0040059918765460

Crabtree, T. R., Alber-Morgan, S. R., & Konrad, M. (2010). The effects of self-monitoring of story elements on the reading comprehension of high school seniors with learning disabilities. *Education and Treatment of Children, 33*, 187–203.

Craft, M. A., Alber, S. R., & Heward, W. L. (1998). Teaching elementary students with developmental disabilities to recruit teacher attention in a general education classroom: Effects on teacher praise and academic productivity. *Journal of Applied Behavior Analysis, 31*, 399–415.

Craig, A. R. (2010). Self-administered behavior modification to reduce nail biting: Incorporating simple technology to ensure treatment integrity. *Behavior Analysis in Practice, 3*(2), 38–41.

Craig, A. R., Nevin, J. A., & Odum, A. L. (2014). Behavioral momentum and resistance to change. In F. K. McSweeney & E. S. Murphy (Eds.), *Wiley-Blackwell handbook of operant and classical conditioning* (pp. 249–274). Oxford, UK: Wiley-Blackwell.

Critchfield, T. S. (1999). An unexpected effect of recording frequency in reactive self-monitoring. *Journal of Applied Behavior Analysis, 32*, 389–391.

Critchfield, T. S. (2011a). To a young basic scientist, about to embark on a program of translational research. *The Behavior Analyst, 34*, 133–148.

Critchfield, T. S. (2011b). Translational contributions of the experimental analysis of behavior. *The Behavior Analyst, 34*, 3–17.

Critchfield, T. S. (2014). Online equivalence-based instruction about statistical inference using written explanation instead of match-to-sample training. *Journal of Applied Behavior Analysis, 47*, 606–611.

Critchfield, T. S. (2014). Skeptic's corner: Punishment—Destructive force or valuable social "adhesive." *Behavior Analysis in Practice, 7*, 36–44.

Critchfield, T. S. (2015). PBIS may not count as classical Applied Behavior Analysis: So what? *Behavior Analysis in Practice, 8*, 99–100.

Critchfield, T. S. (2016). Requiem for the Dead Man test? *The Behavior Analyst*. doi:10.1007/s40614-016-0082-5

Critchfield, T. S. (2018). Efficiency is everything: Promoting efficient practice by harnessing derived stimulus relations. *Behavior Analysis in Practice, 11*, 206–210.

Critchfield, T. S., & Fienup, D. M. (2008). Stimulus equivalence. In S. F. Davis & W. F. Buskist (Eds.), *21st century psychology* (pp. 360–372). Thousand Oaks, CA: Sage.

Critchfield, T. S., & Fienup, D. M. (2010). Using stimulus equivalence technology to teach statistical inference in a group setting. *Journal of Applied Behavior Analysis, 43*, 763–768.

Critchfield, T. S., & Kollins, S. H. (2001). Temporal discounting: Basic research and the analysis of socially important behavior. *Journal of Applied Behavior Analysis, 34*, 101–122.

Critchfield, T. S., & Lattal, K. A. (1993). Acquisition of a spatially defined operant with delayed reinforcement. *Journal of the Experimental Analysis of Behavior, 59*, 373–387.

Critchfield, T. S., & Reed, D. D. (2017). The fuzzy concept of applied behavior analysis research. *The Behavior Analyst, 40*, 123–159. doi:10.1007/s40614-017-0093-x

Critchfield, T. S., & Twyman, J. S. (2014). Prospective instructional design: Establishing conditions for emergent learning. *Journal of Cognitive Education and Psychology, 13*, 201–217.

Critchfield, T. S., Barnes-Holmes, D., & Dougher, M. J. (2018). What Sidman Did—Historical and contemporary significance of research on derived stimulus relations. *Perspectives on Behavior Science, 41*, 9–32.

Critchfield, T. S., Tucker, J. A., & Vuchinich, R. E. (1998). Self-report methods. In K. A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 435–470). New York: Plenum.

Cromwell, O. (1650, August 3). Letter to the general assembly of the Church of Scotland. Available online: http://en.wikiquote.org/wiki/Oliver_Cromwell

Crowley-Koch, B. J., & Van Houten, R. (2013). Automated measurement in applied behavior analysis: A review. *Behavioral Interventions, 28*, 225–240. doi.org/10.1002/bin.1366

Crowley-Koch, B. J., Van Houten, R., & Lim, E. (2011). Effects of pedestrian prompts on motorist yielding at crosswalks. *Journal of Applied Behavior Analysis, 44*, 121–126. doi:10.1901/jaba.2011.44-121

Crutchfield, S. A., Mason, R. A., Chambers, A., Wills, H. P., & Mason, B. A. (2015). Use of a self-monitoring application to reduce stereotypic behavior in adolescents with autism: A preliminary investigation of I-Connect. *Journal of Autism and Developmental Disorders, 45*, 1146–1155.

Cullen, J. M., & Alber-Morgan, S. R. (2015). Technology mediated self-prompting of daily living skills for adolescents and adults with disabilities: A review of the literature. *Education and Training in Autism and Developmental Disabilities, 50*, 43–55.

Cullen, J. M., Alber-Morgan, S. R., Schnell, S. T., & Wheaton, J. E. (2014). Improving reading skills of students with disabilities using Headsprout Comprehension. *Remedial and Special Education, 35*(6), 356–365. doi:10.1177/0741932514534075

Cullen, J. M., Alber-Morgan, S. R., Simmons-Reeds, E. A., & Izzo, M. V. (2017). Effects of self-directed video prompting using iPads on the vocational task completion of young adults with intellectual and developmental disabilities. *Journal of Vocational Rehabilitation, 46*, 361–365. doi:10.3233/JVR-170873

Cummings, A. R., & Carr, J. E. (2005). Functional analysis and treatment of joint dislocation associated with hypermobility syndrome: A single-case analysis. *Journal of Developmental and Physical Disabilities, 17*, 225–236. doi:10.1007/s10882-005-4379-7

Curtis, R. C. (1989). *Self-defeating behaviors: Experimental research, clinical impressions, and practical implications*. New York: Plenum.

Cushing, L. S., & Kennedy, C. H. (1997). Academic effects of providing peer support in general education classrooms on students without disabilities. *Journal of Applied Behavior Analysis, 30*, 139–151.

Cuvo, A. J., & Riva, M. T. (1980). Generalization and transfer between comprehension and production: A comparison of retarded and nonretarded persons. *Journal of Applied Behavior Analysis, 13*, 315–331.

Cuvo, A. J., Lerch, L. J., Leurquin, D. A., Gaffaney, T. J., & Poppen, R. L. (1998). Response allocation to concurrent fixed-ratio reinforcement schedules with work requirements by adults with mental retardation and typical preschool children. *Journal of Applied Behavior Analysis, 31*, 43–63.

da Silva, W. R., de Souza, D. G., & de Rose, J. C. (2006). Relational learning in children with cochlear implants. *Experimental Analysis of Human Behavior Bulletin, 24*, 1–8.

Dachman, R. S., Alessi, G. J., Vrazco, G. J., Fuqua, R. W., & Kerr, R. H. (1986). Development and evaluation of an infant-care training program with first-time fathers. *Journal of Applied Behavior Analysis, 19*, 221–230.

Dack, C., McHugh, L., & Reed, P. (2009). Generalization of causal efficacy judgments after evaluative learning. *Learning & Behavior, 37*, 336–348.

Dallery, J., Meredith, S., Jarvis, B., & Nuzzo, P. A. (2015). Internet-based group contingency management to promote smoking abstinence. *Experimental Clinical Psychopharmacology, 23*(3), 176–183.

Dallery, J., Raiff, B. R., & Grabinski, M. J. (2013). Internet-based contingency management to promote smoking cessation: A randomized controlled study. *Journal of Applied Behavior Analysis, 46*, 750–764. doi:10.1002/jaba.89

Dallery, J., Raiff, B. R., Kim, S. J., Marsch, L. A., Stitzer, M., & Grabinski, M. J. (2017). Nationwide access to an Internet-based contingency management intervention to promote smoking cessation: A randomized controlled trial. *Addiction, 112*(5), 875–883. doi:10.1111/add.13715

Dalton, T., Martella, R., & Marchand-Martella, N. E. (1999). The effects of a self-management program in reducing off-task behavior. *Journal of Behavioral Education, 9*, 157–176.

Daly, P. M., & Ranalli, P. (2003). Using countoons to teach self-monitoring skills. *Teaching Exceptional Children, 35*(5), 30–35.

D'Amato, M. R., Salmon, D. P., Loukas, E., & Tomie, A. (1985). Symmetry and transitivity of conditional relations in monkeys (*Cebus apella*) and pigeons (*Columba livia*). *Journal of the Experimental Analysis of Behavior, 44*, 35–47.

Dams, P.-C. (2012). Case study: A little night music. In R. W. Malott, *I'll stop procrastinating when I get around to it: Plus other cool ways to succeed in school and life using behavior analysis to get your act together* (pp. 6–71, 6–72). Kalamazoo, MI: Behaviordelia.

Daniels, A., & Lattal, A. D. (2009). *Life's a PIC/NIC . . . when you understand behavior*. Cornwall on Hudson, NY: Sloan.

Dardano, J. F. (1973). Self-imposed timeouts under increasing response requirements. *Journal of the Experimental Analysis of Behavior, 19*, 269–287.

Dardig, J. C., & Heward, W. L. (1981). A systematic procedure for prioritizing IEP goals. *The Directive Teacher, 3*, 6–8.

Dardig, J. C., & Heward, W. L. (2016). *Sign here: A contracting book for children and their parents* (2nd ed.). Bucharest, Romania: ACTA. [Also available in Czech, Italian, Polish, Russian, and Turkish language editions.]

Dart, E. H., & Radley, K. C. (2017). The impact of ordinate scaling on the visual analysis of single-case data. *Journal of School Psychology, 63*, 105–118. doi.org/10.1016/j.jsp.2017.03.008

Davenport, C. A., Alber-Morgan, S. R., Clancy, S. M., & Kranak, M. P. (2017). Effects of a picture racetrack game on the expressive vocabulary of deaf preschoolers. *Journal of Deaf Studies and Deaf Education, 2017*, 326–335. doi:10.1093/deafed/enx015

Davis, C., Patte, K., Curtis, C., & Reid, C. (2010). Immediate pleasures and future consequences: A neuropsychological study of binge eating and obesity. *Appetite, 54*, 208–213. doi:10.1016/j.appet.2009.11.002

Davis, C. A., & Reichle, J. (1996). Variant and invariant high-probability requests: Increasing appropriate behaviors in children with emotional-behavioral disorders. *Journal of Applied Behavior Analysis, 29*, 471–482.

Davis, L. L., & O'Neill, R. E. (2004). Use of response cards with a group of students with learning disabilities including those for whom English is a second language. *Journal of Applied Behavior Analysis, 37*, 219–222.

Davis, P. K., & Chittum, R. (1994). A group-oriented contingency to increase leisure activities of adults with traumatic brain injury. *Journal of Applied Behavior Analysis, 27*, 553–554.

Davis-Temple, J., Jung, S., & Sainato, D. M. (2014). Teaching young children with special needs and their peers to play board games: Effects of a least to most prompting procedure to increase independent performance. *Behavior Analysis in Practice, 7*(1), 21–30.

Dawson, J. E., Piazza, C. C., Sevin, B. M., Gulotta, C. S., Lerman, D., & Kelley, M. L. (2003). Use of the high-probability instructional sequence and escape extinction in a child with food refusal. *Journal of Applied Behavior Analysis, 36*, 105–108. doi:10.1901/jaba.2003.36-105

Day, R. M., Rea, J. A., Schussler, N. G., Larsen, S. E., & Johnson, W. L. (1988). A functionally based approach to the treatment of self-injurious behavior. *Behavior Modification, 12*, 565–589.

De Luca, R. B., & Holborn, S. W. (1990). Effects of fixed-interval and fixed-ratio schedules of token reinforcement on exercise with obese and nonobese boys. *Psychological Record, 40*, 67–82.

De Luca, R. B., & Holborn, S. W. (1992). Effects of a variable-ratio reinforcement schedule with changing criteria on exercise in obese and nonobese boys. *Journal of Applied Behavior Analysis, 25*, 671–679.

De Martini-Scully, D., Bray, M. A., & Kehle, T. J. (2000). A packaged intervention to reduce disruptive behaviors in general education students. *Psychology in the Schools, 37*(2), 149–156.

de Melo, C. M., de Castro, M. S. L. B., & de Rose, J. C. (2015). Some relations between ethics, culture, and technology in B. F. Skinner. *Behavior and Social Issues, 24*, 39–55.

de Rose, J. C., de Souza, D. G., & Hanna, E. S. (1996). Teaching reading and spelling: Exclusion and stimulus equivalence. *Journal of Applied Behavior Analysis, 29*, 451–469.

de Rose, J. C., de Souza, D. G., Rossito, A. L., & de Rose, T. M. S. (1992). Stimulus equivalence and generalization in reading after matching-to-sample by exclusion. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 69–82). Reno, NV: Context Press.

de Souza, D. G., de Rose, J. C., Faleiros, T. C., Bortoloti, R., Hanna, E. S., & McIlvane, W. J. (2009). Teaching generative reading via recombination of minimal textual units: A legacy of verbal behavior to children in Brazil. *International Journal of Psychology and Psychological Therapy, 9*, 19–44.

de Souza Barba, L. (2012). Operant variability: A conceptual analysis. *The Behavior Analyst, 35*, 213–227.

de Zubicaray, G., & Clair, A. (1998). An evaluation of differential reinforcement of other behavior, differential reinforcement of incompatible behaviors, and restitution for the management of aggressive behaviors. *Behavioral Interventions, 13*, 157–168.

Deaver, C. M., Miltenberger, R. G., & Stricker, J. M. (2001). Functional analysis and treatment of hair twirling in a young child. *Journal of Applied Behavior Analysis, 34*, 535–538.

Debert, P., Matos, A. M., & McIlvane, W. J. (2007). Conditional relations with compound abstract stimuli using a go/no-go procedure. *Journal of the Experimental Analysis of Behavior, 87*, 89–96.

DeCatanzaro, D., & Baldwin, G. (1978). Effective treatment of self-injurious behavior through a forced arm exercise. *Journal of Applied Behavior Analysis, 1*, 433–439.

DeFulio, A. (2011). Using words instead of jumbled characters as stimuli in keyboard training facilitates fluent performance. *Journal of Applied Behavior Analysis, 44*, 921–924.

DeGrandpre, R. J., Bickel, W. K., & Higgins, S. T. (1992). Emergent equivalence relations between interoceptive (drug) and exteroceptive (visual) stimuli. *Journal of the Experimental Analysis of Behavior, 58*, 9–18.

Deitz, D. E. D. (1977). An analysis of programming DRL schedules in educational settings. *Behavior Research and Therapy, 15*, 103–111.

Deitz, S. M. (1982). Defining applied behavior analysis: An historical analogy. *The Behavior Analyst, 5*, 53–64.

Deitz, S. M., & Repp, A. C. (1973). Decreasing classroom misbehavior through the use of DRL schedules of reinforcement. *Journal of Applied Behavior Analysis, 6*, 457–463.

Deitz, S. M., & Repp, A. C. (1983). Reducing behavior through reinforcement. *Exceptional Education Quarterly, 3*, 34–46.

Deitz, S. M., Slack, D. J., Schwarzmueller, E. B., Wilander, A. P., Weatherly, T. J., & Hilliard, G. (1978). Reducing inappropriate behavior in special classrooms by reinforcing average interresponse times: Interval DRL. *Behavior Therapy, 9*, 37–46.

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis, 29*, 519–533.

DeLeon, I. G., Anders, B. M., Rodriguez-Catter, V., & Neidert, P. L. (2000). The effects of noncontingent access to single-versus multiple-stimulus sets on self-injurious behavior. *Journal of Applied Behavior Analysis, 33*, 623–626.

DeLeon, I. G., Bullock, C. E., & Catania, A. C. (2013). Arranging reinforcement contingencies in applied settings: Fundamentals and implications of recent basic and applied research. In G. J. Madden, W. V. Dube, T. D. Hackenberg, G. P. Hanley, & K. A. Lattal (Eds.), *APA handbook of behavior analysis, Vol. 2: Translating principles into practice* (pp. 47–75).

DeLeon, I. G., Fisher, W. W., Rodriguez-Catter, V., Maglieri, K., Herman, K., & Marhefka, J.-M. (2001). Examination of relative reinforcement effects of stimuli identified through pretreatment and daily brief preference assessments. *Journal of Applied Behavior Analysis, 34*, 463–473.

DeLeon, I. G., Frank, M. A., Gregory, M. K., & Allman, M. J. (2009). On the correspondence between preference assessment outcomes and progressive-ratio schedule assessments of stimulus value. *Journal of Applied Behavior Analysis, 42*, 729–733.

DeLeon, I. G., Hagopian, L. P., Rodriguez-Catter, V., Bowman, L. G., Long, E. S., & Boelter, E. W. (2008). Increasing wearing of prescription glasses in individuals with mental retardation. *Journal of Applied Behavior Analysis, 41*, 137–142.

DeLeon, I. G., Iwata, B. A., Conners, J., & Wallace, M. D. (1999). Examination of ambiguous stimulus preferences with duration-based measures. *Journal of Applied Behavior Analysis, 32*, 111–114.

DeLeon, I. G., Iwata, B. A., Goh, H. L., & Worsdell, A. S. (1997). Emergence of reinforcer preference as a function of schedule requirements and stimulus similarity. *Journal of Applied Behavior Analysis, 30*, 439–449.

Delgado, J. A. P., & Greer, R. D. (2009). The effects of peer monitoring training on the emergence of the capability to learn from observing instruction received by peers. *The Psychological Record, 59*, 407–434.

Deliperi, P., Vladescu, J. C., Reeve, K. F., Reeve, S. A., & DeBar, R. M. (2015). Training staff to implement a paired-stimulus preference assessment using video modeling with voiceover instruction. *Behavioral Interventions, 30*, 314–332. doi:10.1002/bin.1421

DeLissovoy, V. (1963). Head banging in early childhood: A suggested cause. *Journal of Genetic Psychology, 102*, 109–114.

Della Rosa, K. A., Fellman, D., DeBiase, C., DeQuinzio, J. A., & Taylor, B. A. (2015). The effects of using a conditioned stimulus to cue DRO schedules. *Behavioral Interventions, 30*, 219–230. doi:10.1002/bin.1409

Delli Bovi, G., Vladescu, J. C., DeBar, R. M., Carroll, R. A., & Sarokoff, R. A. (2017). Using video modeling with voice-over instruction to train public school staff to implement a preference assessment. *Behavior Analysis in Practice, 10*(1), 72–76.

Delprato, D. J. (2002). Countercontrol in behavior analysis. *The Behavior Analyst, 25*, 191–200.

Delprato, D. J., & Midgley, B. D. (1992). Some fundamentals of B. F. Skinner's behaviorism. *American Psychologist, 48*, 1507–1520.

DeLuca, R. V., & Holborn, S. W. (1992). Effects of a variable-ratio reinforcement schedule with changing criteria on exercise in obese and nonobese boys. *Journal of Applied Behavior Analysis, 25*, 671–679.

Dempsey, C. M., Iwata, B. A., Fritz, J. N., & Rolider, N. U. (2012). Observer training revisited: A comparison of in vivo and video instruction. *Journal of Applied Behavior Analysis, 45*(4), 827–832.

Denton, R. F., & Meindl, J. N. (2016). The effect of colored overlays on reading fluency in individuals with dyslexia. *Behavior Analysis in Practice, 9*(3), 191–198.

Deochand, N. (2017). Automating phase change lines and their labels using Microsoft Excel(R). *Behavior Analysis in Practice*. Published online: 12 January 2017. doi 10.1007/s40617-016-0169-1

Deochand, N., Costello, M. S., & Fuqua, R. W. (2015). Phase-change lines, scale breaks, and trend lines using Excel 2013. *Journal of Applied Behavior Analysis, 48*, 478–493.

DeProspero, A., & Cohen, S. (1979). Inconsistent visual analyses of intrasubject data. *Journal of Applied Behavior Analysis, 12*, 573–579.

DeQuinzio, J. A., & Taylor, B. A. (2015). Teaching children with autism to discriminate the reinforced and nonreinforced responses of others: Implications for observational learning. *Journal of Applied Behavior Analysis, 48*, 38–51. doi:10.1002/jaba.192

Derby, K. M., Fisher, W. W., & Piazza, C. C. (1996). The effects of contingent and noncontingent attention on self-injury and self-restraint. *Journal of Applied Behavior Analysis, 29*, 107–110.

Derby, K. M., Wacker, D. P., Berg, W., DeRaad, A., Ulrich, S., Asmus, J., . . . Stoner, E. A. (1997). The long-term effects of functional communication training in home settings. *Journal of Applied Behavior Analysis, 30*, 507–531.

Derby, K. M., Wacker, D. P., Sasso, G., Steege, M., Northup, J., Cigrand, K., & Asmus, J. (1992). Brief functional assessment techniques to evaluate aberrant behavior in an outpatient setting: A summary of 79 cases. *Journal of Applied Behavior Analysis, 25*, 713–721. doi:10.1901/jaba.1992.25-713

DeRosa, N. M., Roane, H. S., Bishop, J. R., & Silkowski, E. L. (2016). The combined effects of noncontingent reinforcement and punishment on the reduction of rumination. *Journal of Applied Behavior Analysis, 49*, 680–685.

DeSouza, A. A., Fisher, W. W., & Rodriguez, N. M. (2019). Facilitating the emergence of convergent intraverbals in children with autism. *Journal of Applied Behavior Analysis, 52*, 28–49. doi: 10.1002/jaba.520

Devine, B., Carp, C. L., Hiett, K. A., & Petursdottir, A. I. (2016). Emergence of intraverbal responding following tact instruction with compound stimuli. *The Analysis of Verbal Behavior*, *32*, 154–170. doi.org/10.1007/s40616-016-0062-6

Devine, S. L., Rapp, J. T., Testa, J. R., Henrickson, M. L., & Schnerch, G. (2011). Detecting changes in simulated events using partial-interval recording and momentary time sampling III: Evaluating sensitivity as a function of session length. *Behavioral Interventions*, *24*, 103–124.

DeVries, J. E., Burnette, M. M., & Redmon, W. K. (1991). AIDS prevention: Improving nurses' compliance with glove wearing through performance feedback. *Journal of Applied Behavior Analysis*, *24*, 705–711.

Dewey, J. (1939). *Experience and education*. New York: Macmillan.

Dickerson, E. A., & Creedon, C. F. (1981). Self-selection of standards by children: The relative effectiveness of pupil-selected and teacher-selected standards of performance. *Journal of Applied Behavior Analysis*, *14*, 425–433.

Dickman, S. E., Bright, C. N., Montgomery, D. H., & Miguel, C. F. (2012). The effects of response interruption and redirection (RIRD) and differential reinforcement on vocal stereotypy and appropriate vocalizations. *Behavioral Interventions*, *27*, 185–192. doi:10.1002/bin.1348

Dickson, M. J., & Vargo, K. K. (2017). Training kindergarten students lockdown drill procedures using behavioral skills training. *Journal of Applied Behavior Analysis*, *50*, 407–412. doi:10.1002/jaba.369

Didden, R., Prinsen, H., & Sigafoos, J. (2000). The blocking effect of pictorial prompts on sight-word reading. *Journal of Applied Behavior Analysis*, *33*, 317–320. doi:10.1901/jaba.2000.33-317

Didden, R., Sigafoos, J., O'Reilly, M. F., Lancioni, G. E., & Sturmey, P. (2007). A multisite cross-cultural replication of Upper's (1974) unsuccessful self-treatment of writer's block. *Journal of Applied Behavior Analysis*, *40*, 773.

Dietz, T., Gardner, G. T., Gilligan, J., Stern, P., & Vandenbergh, M. P. (2009). Household actions can provide a behavioral wedge to rapidly reduce US carbon emissions. *Proceeding of the National Academy of Sciences*, *106*(44), 18452–18456.

DiGennaro Reed, F. D., & Codding, R. S. (2014). Advancements in procedural fidelity assessment and intervention: Introduction to the special issue. *Journal of Behavioral Education*, *23*, 1–18. doi 10.1007/s10864-013-9191-3

DiGennaro Reed, F. D., Reed, D. D., Baez, C. N., & Maguire, H. (2011). A parametric analysis of errors of commission during discrete-trial training. *Journal of Applied Behavior Analysis*, *44*, 611–615.

Dillenburger, K., & Keenan, M. (2005). Bereavement: A D.I.S.C. analysis. *Behavior and Social Issues*, *14*, 92–112.

Diller, J. W., & Lattal, K. A. (2008). Radical behaviorism and Buddhism: Complimentarities and conflicts. *The Behavior Analyst*, *31*, 163–177.

Diller, J. W., Barry, R. J., & Gelino, B. W. (2016). Visual analyses of data in a multielement design. *Journal of Applied Behavior Analysis*, *49*, 980–985. doi:10.1002/jaba.325

Dinsmoor, J. A. (1952). A discrimination based on punishment. *Quarterly Journal of Experimental Psychology*, *4*, 27–45.

Dinsmoor, J. A. (1995a). Stimulus control: Part I. *The Behavior Analyst*, *18*, 51–68.

Dinsmoor, J. A. (1995b). Stimulus control: Part II. *The Behavior Analyst*, *18*, 253–269.

Dinsmoor, J. A. (2003). Experimental. *The Behavior Analyst*, *26*, 151–153.

Dipipi-Hoy, C., & Jitendra, A. (2004). A parent-delivered intervention to teach purchasing skills to young adults with disabilities. *The Journal of Special Education*, *38*, 144–157.

Dittlinger, L. H., & Lerman, D. C. (2011). Further analysis of picture interference when teaching word recognition to children with autism. *Journal of Applied Behavior Analysis*, *44*, 341–349. doi:10.1901/jaba.2011.44-341

Dixon, M. R. (2014). *The PEAK relational training system: Direct training module*. Carbondale, IL: Shawnee Scientific Press.

Dixon, M. R. (2016). *The PEAK relational training system: Transformation module*. Carbondale, IL: Shawnee Scientific.

Dixon, M. R., & Cummings, A. (2001). Self-control in children with autism: Response allocation during delays to reinforcement. *Journal of Applied Behavior Analysis*, *34*, 491–495. doi:10.1901/jaba.2001.

Dixon, M. R., & Falcomata, T. S. (2004). Preference for progressive delays and concurrent physical therapy exercise in an adult with acquired brain injury. *Journal of Applied Behavior Analysis*, *37*, 101–105.

Dixon, M. R., & Holcomb, S. (2000). Teaching self-control to small groups of dually diagnosed adults. *Journal of Applied Behavior Analysis*, *33*, 611–614. doi:10.1901/jaba.2000.33-611

Dixon, M. R., & Holton, B. (2009). Altering the magnitude of delay discounting by pathological gamblers. *Journal of Applied Behavior Analysis*, *42*, 269–275.

Dixon, M. R., Baker, J. C., & Sadowski, K. A. (2011). Applying Skinner's analysis of verbal behavior to persons with dementia. *Behavior Therapy*, *42*, 120–126. doi:10.1016/j.beth.2010.05.002

Dixon, M. R., Belisle, J., McKeel, A., Whiting, S., Speelman, R., Daar, J. H., & Rowsey, K. (2017). An internal and critical review of the PEAK relational training system for children with autism and related intellectual disabilities: 2014–2017. *The Behavior Analyst*, *40*, 493–521.

Dixon, M. R., Belisle, J., Rehfeldt, R. A., & Root, W. B. (2018). Why we are still not acting to save the world: The upward challenge of a post-Skinnerian behavior science. *Perspectives on Behavior Science*, *41*, 241–267.

Dixon, M. R., Benedict, H., & Larson, T. (2001). Functional analysis and treatment of inappropriate verbal behavior. *Journal of Applied Behavior Analysis*, *34*, 361–363.

Dixon, M. R., Guercio, J., Falcomata, T., Horner, M. J., Root, S., Newell, C., & Zlomke, K. (2004). Exploring the utility of functional analysis methodology to assess and treat problematic verbal behavior in persons with acquired brain injury. *Behavioral Interventions*, *19*, 91–102.

Dixon, M. R., Hayes, L. J., Binder, L. M., Manthey, S., Sigman, C., & Zdanowski, D. M. (1998). Using a self-control training procedure to increase appropriate behavior. *Journal of Applied Behavior Analysis*, *31*, 203–210. doi:10.1901/jaba.1998.31-203

Dixon, M. R., Jackson, J. W., Small, S. L., Horner-King, M. J., Mui Ker Lik, N., Garcia, Y., & Rosales, R. (2009). Creating single-subject design graphs in Microsoft Excel 2007. *Journal of Applied Behavior Analysis*, *42*, 277–293.

Dixon, M. R., Jacobs, E. A., & Sanders, S. (2006). Contextually controlled delay discounting of pathological gamblers. *Journal of Applied Behavior Analysis*, *39*, 413–422.

Dixon, M. R., Marley, J., & Jacobs, E. A. (2003). Delay discounting by pathological gamblers. *Journal of Applied Behavior Analysis*, *36*, 449–458.

Dixon, M. R., Rehfeldt, R., & Randich, L. (2003). Enhancing tolerance to delayed reinforcers: The role of intervening activities. *Journal of Applied Behavior Analysis*, *36*, 263–266. doi:10.1901/jaba.2003.36-263

Dixon, M. R., Whiting, S. W., Gunnarsson, K. F., Daar, J. H., & Rowsey, K. E. (2015). Trends in behavior-analytics gambling research and treatment. *The Behavior Analyst*, *38*, 179–202. doi:10.1007/s40614-015-0027-4

Dogoe, M., & Banda, D. R. (2009). Review of research using constant time delay to teach chained tasks to persons with developmental disabilities. *Education and Training in Autism and Developmental Disabilities*, *44*, 177–186. doi.org/10.1002/jaba.387

Doke, L. A., & Risley, T. R. (1972). The organization of day care environments: Required vs. optional activities. *Journal of Applied Behavior Analysis*, *5*, 453–454.

Doll, C., McLaughlin, T. F., & Barretto, A. (2013). The token economy: A review and evaluation. *International Journal of Basic and Applied Science*, *2*(1), 131–149.

Domjan, M. (2016). Elicited versus emitted behavior: Time to abandon the distinction. *Journal of the Experimental Analysis of Behavior*, *105*, 231–245.

Donahoe, J. W., & Palmer, D. C. (1994). *Learning and complex behavior*. Boston: Allyn and Bacon.

Donaldson, J. M., & Normand, M. P. (2009). Using goal setting, self-monitoring, and feedback to increase calorie expenditure in obese adults. *Behavioral Interventions*, *24*, 73–83. http://dx.doi.org/10.1002/bin.277

Donaldson, J. M., & Vollmer, T. R. (2011). An evaluation and comparison of time-out procedures with and without release contingencies. *Journal of Applied Behavior Analysis*, *44*, 693–705.

Donaldson, J. M., DeLeon, I. G., Fisher, A. B., & Kahng, S. W. (2014). Effects of and preference for conditions of token earn versus token loss. *Journal of Applied Behavior Analysis*, *47*, 537–548.

Donaldson, J. M., Fisher, A. B., & Kahng, S. (2017). Effects of the Good Behavior Game on individual student behavior. *Behavior Analysis: Research and Practice*, *17*, 207–216.

Donaldson, J. M., Vollmer, T. R., Krous, T., Downs, S., & Berard, K. P. (2011). An evaluation of the Good Behavior Game in kindergarten classrooms. *Journal of Applied Behavior Analysis*, *44*, 605–609.

Dorigo, M., & Colombetti, M. (1998). *Robot shaping: An experiment in behavior engineering*. Cambridge, MA: MIT Press.

Dorow, L. G., & Boyle, M. E. (1998). Instructor feedback for college writing assignments in introductory classes. *Journal of Behavioral Education*, *8*, 115–129.

Dougher, M. (2009). In Memoriam: Sidney W. Bijou, 1908–2009. *The Behavior Analyst*, *32*, 363–364.

Dougher, M. J., Augustson, E. M., Markham, M. R., Greenway, D. E., & Wulfert, E. (1994). The transfer of respondent eliciting and extinction functions through stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior*, *62*, 331–351.

Dougher, M. J., Hamilton, D. A., Fink, B. C. & Harrington, J. (2007). Transformation of the discriminative and eliciting functions of generalized relational stimuli. *Journal of the Experimental Analysis of Behavior*, *88*, 179–197.

Doughty, A. H., & Hopkins, M. N. (2011). Reducing stimulus overselectivity through an increased observing-response requirement. *Journal of Applied Behavior Analysis*, *44*, 653–657. doi:10.1901/jaba.2011.44-653

Downing, J. A. (1990). Contingency contracting: A step-by-step format. *Teaching Exceptional Children*, *26*(2), 111–113.

Downs, H. E., Miltenberger, R., Biedronski, J., & Witherspoon, L. (2015). The effects of video self-evaluation on skill acquisition with yoga postures. *Journal of Applied Behavior Analysis*, *48*, 930–935. doi:10.1002/jaba.248

Dowrick, P. W., & Dove, C. (1980). The use of self-monitoring to improve the swimming performance of spina bifida children. *Journal of Applied Behavior Analysis*, *13*, 51–56.

Dozier, C. L., Iwata, B. A., Thomason-Sassi, J., Worsdell, A. S., & Wilson, D. M. (2012). A comparison of two pairing procedures to establish praise as a reinforcer. *Journal of Applied Behavior Analysis*, *45*, 721–735.

Drabman, R. S., Hammer, D., & Rosenbaum, M. S. (1979). Assessing generalization in behavior modification with children: The generalization map. *Behavioral Assessment*, *1*, 203–219.

Drabman, R. S., Spitalnik, R., & O'Leary, K. D. (1973). Teaching self-control to disruptive children. *Journal of Abnormal Psychology*, *82*, 10–16.

Drasgow, E., Halle, J., & Ostrosky, M. M. (1998). Effects of differential reinforcement on the generalization of a replacement mand in three children with severe language delays. *Journal of Applied Behavior Analysis*, *31*, 357–374.

Dube, W. V., & McIlvane, W. J. (1996). Some implications of a stimulus control topography analysis for emergent stimulus classes. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 197–218). Amsterdam, Netherlands: Elsevier.

Dube, W. V., & McIlvane, W. J. (1999). Reduction of stimulus overselectivity with nonverbal differential observing responses. *Journal of Applied Behavior Analysis*, *32*, 25–33. doi:10.1901/jaba.1999.32-25

Dube, W. V., Farber, R. S., Mueller, M. R., Grant, E., Lorin, L., & Deutsch, C. K. (2016). Stimulus overselectivity in autism, Down syndrome, and typical development. *American Journal on Intellectual and Developmental Disabilities*, *121*, 219–235. doi:10.1352/1944-7558-121.3.219

Dube, W. V., McIlvane, W. J., Mackay, H. A., & Stoddard, L. T. (1987). Stimulus class membership established via stimulus-reinforcer relations. *Journal of the Experimental Analysis of Behavior*, *47*, 159–175.

Dube, W. V., McIlvane, W. J., Maguire, R. W., Mackay, H. A., & Stoddard, L. T. (1989). Stimulus class formation and stimulus-reinforcer relations. *Journal of the Experimental Analysis of Behavior*, *47*, 159–175.

Duchaine, E. L., Jolivette, K., & Fredrick, L. D. (2011). The effect of teacher coaching with performance feedback on behavior-specific praise in inclusion classrooms. *Education & Treatment of Children*, *34*, 209–227. http://dx.doi.org/10.1353/etc.2011.0009

Ducharme, J. M., & Rushford, N. (2001). Proximal and distal effects of play on child compliance with a brain-injured parent. *Journal of Applied Behavior Analysis*, *34*, 221–224.

Dugdale, N. A., & Lowe, C. F. (1990). Naming and stimulus equivalence. In D. E. Blackman & H. Lejeune (Eds.), *Behaviour analysis in theory and practice: Contributions and controversies* (pp. 115–138). Hove, UK: Erlbaum.

Duker, P. C., & Seys, D. M. (1996). Long-term use of electrical aversion treatment with self-injurious behavior. *Research in Developmental Disabilities*, *17*, 293–301.

Duker, P. C., & van Lent, C. (1991). Inducing variability in communicative gestures used by severely retarded individuals. *Journal of Applied Behavior Analysis*, *24*, 379–386.

Dunkel-Jackson, S. M., Dixon, M. R., & Szekely, S. (2016). Self-control as generalized operant behavior by adults with autism spectrum disorder. *Journal of Applied Behavior Analysis*, *49*, 705–710. doi:10.1002/jaba.315

Dunlap, G., & Johnson, J. (1985). Increasing the independent responding of autistic children with unpredictable supervision. *Journal of Applied Behavior Analysis*, *18*, 227–236.

Dunlap, G., Kern-Dunlap, L., Clarke, S., & Robbins, F. R. (1991). Functional assessment, curricular revision, and severe behavior problems. *Journal of Applied Behavior Analysis*, *24*, 387–397.

Dunlap, G., Koegel, R. L., Johnson, J., & O'Neill, R. E. (1987). Maintaining performance of autistic clients in community settings with delayed contingencies. *Journal of Applied Behavior Analysis*, *20*, 185–191.

Dunlap, L. K., & Dunlap, G. (1989). A self-monitoring package for teaching subtraction with regrouping to students with learning disabilities. *Journal of Applied Behavior Analysis*, *22*, 309–314.

Dunn, L. M., & Dunn, D. M. (1997). *Peabody picture vocabulary test* (3rd ed.). Circle Pines, MN: American Guidance Service, Inc.

Dunne, S., Foody, M., Barnes-Holmes, Y., Barnes-Holmes, D., & Murphy, C. (2014). Facilitating repertoires of coordination, opposition distinction, and comparison in young children with autism. *Behavioral Development Bulletin*, *19*, 37–47.

Dupuis, D. L., Lerman, D. C., Tsami, L., & Shireman, M. L. (2015). Reduction of aggression evoked by sounds using noncontingent reinforcement and timeout. *Journal of Applied Behavior Analysis*, *48*, 669–674.

Durand, V. M. (1999). Functional communication training using assistive devices: Recruiting natural communities of reinforcement. *Journal of Applied Behavior Analysis*, *32*, 247–267.

Durand, V. M., & Carr, E. G. (1987). Social influences on "self-stimulatory" behavior: Analysis and treatment. *Journal of Applied Behavior Analysis*, *20*, 119–132.

Durand, V. M., & Carr, E. G. (1992). An analysis of maintenance following functional communication training. *Journal of Applied Behavior Analysis*, *25*, 777–794.

Durand, V. M., & Crimmins, D. (1992). *The Motivation Assessment Scale*. Topeka, KS: Monaco & Associates.

Durlach, P. J., Elliman, N. A., & Rogers, P. J. (2002). Drinking while thirsty can lead to conditioned increases in consumption. *Appetite*, *39*, 119–125.

Duvinsky, J. D., & Poppen, R. (1982). Human performance on conjunctive fixed-interval fixed-ratio schedules. *Journal of the Experimental Analysis of Behavior*, *37*, 243–250.

Dwyer-Moore, K. J., & Dixon, M. R. (2007). Functional analysis and treatment of problem behavior of elderly adults in long-term care. *Journal of Applied Behavior Analysis*, *40*, 679–683.

Dyer, K., Dunlap, G., & Winterling, V. (1990). Effects of choice making on the serious problem behaviors of students with severe handicaps. *Journal of Applied Behavior Analysis*, *23*, 515–524.

Dymond, S. (2009). A contemporary behavior analysis of anxiety and avoidance. *The Behavior Analyst*, *32*, 7–27.

Dymond, S., & Barnes, D. (1994). A transfer of self-discrimination response functions through equivalence relations. *Journal of the Experimental Analysis of Behavior*, *62*, 251–267.

Dymond, S., & Barnes, D. (1995). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness, more than, and less than. *Journal of the Experimental Analysis of Behavior*, *64*, 163–184.

Dymond, S., & Barnes, D. (1996). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness and opposition. *The Psychological Record*, *46*, 271–300.

Dymond, S., & Rehfeldt, R. A. (2000). Understanding complex behavior: The transformation of stimulus functions. *The Behavior Analyst*, *23*, 239–254.

Dymond, S., & Roche, B. (2009). A contemporary behavior analysis of anxiety and avoidance. *The Behavior Analyst*, *32*, 7–27.

Dymond, S., O'Hora, D., Whelan, R., & O'Donovan, A. (2006). Citation analysis of Skinner's *Verbal Behavior*: 1984–2004. *The Behavior Analyst*, *29*, 75–88. doi.org/10.1007/BF03392118

Dymond, S., Schlund, M. W., Roche, B., & Whelan, R. (2014). The spread of fear: Symbolic generalization mediates graded threat-avoidance in specific phobia. *Quarterly Journal of Experimental Psychology, 67*, 247–259.

Ebanks, M. E., & Fisher, W. W. (2003). Altering the timing of academic prompts to treat destructive behavior maintained by escape. *Journal of Applied Behavior Analysis, 36*, 355–359.

Eckert, T. L., Ardoin, S. P., Daly, E. J., III, & Martens, B. K. (2002). Improving oral reading fluency: A brief experimental analysis of combining an antecedent intervention with consequences. *Journal of Applied Behavior Analysis, 35*, 271–281.

Ecott, C. L., Foate, B. A. L., Taylor, B., & Critchfield, T. S. (1999). Further evaluation of reinforcer magnitude effects in noncontingent schedules. *Journal of Applied Behavior Analysis, 32*, 529–532.

Edgerton, L., & Wine, B. (2017). Speak up: Increasing conversational volume in a child with autism spectrum disorder. *Behavior Analysis in Practice, 10*, 407–411. doi.org/10.1007/s40617-016-0168-2

Edgerton, L., & Wine, B. (2017). Speak up: Increasing conversational volume in a child with autism spectrum disorder. *Behavior Analysis in Practice, 10*(4), 407–410. doi 10.1007/s40617-016-0168-2.

Edwards, K. J., & Christophersen, E. R. (1993). Automated data acquisition through time-lapse videotape recording. *Journal of Applied Behavior Analysis, 24*, 503–504.

Edwards, T. L., Cox, C., Weetjens, B., Tewelde, T., & Poling, A. (2015). Giant African pouched rats (*Cricetomys gambianus*) that work on tilled soil accurately detect land mines. *Journal of Applied Behavior Analysis, 48*, 696–700.

Edwards, T. L., La Londe, K. B., Cox, C., Weetjens, B., & Poling, A. (2016). Effects of schedules of reinforcement on pouched rats' performance in urban search-and-rescue training. *Journal of Applied Behavior Analysis, 49*, 199–204.

Egel, A. L. (1981). Reinforcer variation: Implications for motivating developmentally disabled children. *Journal of Applied Behavior Analysis, 14*, 345–350.

Eikeseth, S. (2009). Outcome of comprehensive psycho-educational interventions for young children with autism. *Research Developmental Disabilities, 30*(1), 158–178.

Eikeseth, S., & Smith, D. P. (2013). An analysis of verbal stimulus control in intraverbal behavior: Implications for practice and applied research. *The Analysis of Verbal Behavior, 29*, 125–135.

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *American Journal on Intellectual and Developmental Disabilities, 115*, 364–380.

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2010). Using participant data to extend the evidence base for intensive behavioral intervention for children with autism. *Journal of Clinical Child & Adolescent Psychology, 38*(3), 439–450. doi:10.1080/15374410902851739

Elias, N. C., Goyos, C., Saunders, M., & Saunders, R. (2008). Teaching manual signs to adults with mental retardation using matching-to-sample procedures with stimulus equivalence. *The Analysis of Verbal Behavior, 24*, 1–13.

Elliot, S. N., & Von Brock Treuting, M. (1991). The Behavior Intervention Rating Scale: Development and validation of a pretreatment acceptability and effectiveness measure. *Journal of School Psychology, 29*, 43–51.

Elliot, S. N., Busse, R. T., & Shapiro, E. S. (1999). Intervention techniques for academic problems. In C. R. Reynolds and T. B. Gutkin (Eds.), *The handbook of school psychology* (3rd ed.) (pp. 664–685). New York: John Wiley & Sons.

Embry, D. D. (2002). The Good Behavior Game: A best practice candidate as a universal behavioral vaccine. *Clinical Children Family Psychology Review, 5*(4), 273–297.

Embry, D. D., & Biglan, A. (2008). Evidence-based kernels: Fundamental units of behavioral influence. *Clinical Child and Family Psychology Review, 11*, 75–113.

Endicott, K., & Higbee, T. S. (2007). Contriving motivating operations to evoke mands for information in preschoolers with autism. *Research in Autism Spectrum Disorders, 1*, 210–217.

Engelman, K., Altus, D., & Mathews, M. (1999). Increasing engagement in daily activities by older adults with dementia. *Journal of Applied Behavior Analysis, 32*, 107–110.

Engelman, K. K., Altus, D. E., Mosier, M. C., & Mathews, R. M. (2003). Brief training to promote the use of less intrusive prompts by nursing assistants in a dementia care unit. *Journal of Applied Behavior Analysis, 36*, 129–132. http://dx.doi.org/10.1901/jaba.2003.36-129

Engelmann, S. (1975). *Your child can succeed*. New York: Simon & Shuster.

Engelmann, S., & Carnine, D. (1982). *Theory of instructional design: Principles and applications*. New York: Irvington Press.

Engelmann, S., & Colvin, D. (1983). *Generalized compliance training: A direct-instruction program for managing severe behavior problems*. Austin, TX: PRO-ED.

Epling, W. F., & Pierce, W. D. (1983). Applied behavior analysis: New directions from the laboratory. *The Behavior Analyst, 6*, 27–37.

Epstein, R. (1982). *Skinner for the classroom*. Champaign, IL: Research Press.

Epstein, R. (1990). Generativity theory and creativity. In M. A. Runco & R. S. Albert (Eds.), *Theories of creativity* (pp. 116–140). Newbury Park, CA: Sage.

Epstein, R. (1991). Skinner, creativity, and the problem of spontaneous behavior. *Psychological Science, 2*, 362–370.

Epstein, R. (1996). *Cognition, creativity, and behavior: Selected essays*. Westport, CT: Praeger.

Epstein, R. (1997). Skinner as self-manager. *Journal of Applied Behavior Analysis, 30*, 545–568.

Epstein, R. (2012). Crying babies. *Journal of Contextual Behavioral Science, 1*, 43–48.

Epstein, R., Mejia, J., & Robertson, R. E. (2016). The frequency profile: An informative method for graphing the behavior of individuals post hoc or in real time. *Behavior Analysis in Practice*. doi.org/10.1037/bar0000052

Erford, B. T. (1999). A modified time-out procedure for children with noncompliant or defiant behaviors. *Professional School Counseling, 2*(3), 205–210.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49*, 725–747.

Ervin, R., Radford, P., Bertsch, K., Piper, A., Ehrhardt, K., & Poling, A. (2001). A descriptive analysis and critique of the empirical literature on school-based functional assessment. *School Psychology Review, 30*, 193–210.

Esch, B. E. (2014). The early echoic skills assessment. In M. L. Sundberg (Ed.), *The verbal behavior milestones assessment and placement program: The VB-MAPP* (2nd ed.) (pp. 51–52). Concord, CA: AVB Press.

Esch, B. E., Carr, J. E., & Grow, L. L. (2009). Evaluation of an enhanced stimulus–stimulus pairing procedure to increase early vocalizations of children with autism. *Journal of Applied Behavior Analysis, 42*, 225–241. doi:10.1901/jaba.2009.42-225

Esch, B. E., LaLonde, K. B., & Esch, J. W. (2010). Speech and language assessment: A verbal behavior analysis. *The Journal of Speech-Language Pathology and Applied Behavior Analysis, 5*, 166–191.

Esch, J. W., Esch, B. E., & Love, J. R. (2009). Increasing vocal variability in children with autism using a lag schedule of reinforcement. *The Analysis of Verbal Behavior, 25*, 73–78.

Esch, J. W., Esch, B. E., McCart, J. D., & Petursdottir, A. I. (2010). An assessment of self-echoic behavior in young children. *The Analysis of Verbal Behavior, 26*, 3–13. doi.org/10.1007/BF03393078

Estes, W. K., & Skinner, B. F. (1941). Some quantitative properties of anxiety. *Journal of Experimental Psychology, 29*, 390–400.

Ewry, D. M., & Fryling, M. J. (2016). Evaluating the high-probability instructional sequence to increase the acceptance of foods with an adolescent with autism. *Behavior Analysis in Practice, 9*, 380–383. doi:10.1007/s40617-015-0098-4

Fahmie, T. A., & Hanley, G. P. (2008). Progressing toward data intimacy: A review of within-session data analysis. *Journal of Applied Behavior Analysis 41*, 319–331. doi:10.1901/jaba.2008.41-319

Fahmie, T. A., Iwata, B. A., & Jann, K. E. (2015). Comparison of edible and leisure reinforcers. *Journal of Applied Behavior Analysis, 48*, 331–343. doi:10.1002/jaba.200

Fahmie, T. A., Iwata, B. A., & Mead, S. C. (2016). Within-subject analysis of a prevention strategy for problem behavior. *Journal of Applied Behavior Analysis, 49*, 915–926. doi:10.1002/jaba.343

Falcomata, T. S., Roane, H. S., Hovanetz, A. N., Kettering, T. L., & Keeney, K. M. (2004). An evaluation of response cost in the treatment of inappropriate vocalizations maintained by automatic reinforcement. *Journal of Applied Behavior Analysis, 37*, 83–87.

Falk, J. L. (1961). Production of polydipsia in normal rats by an intermittent food schedule. *Science*, *133*, 195–196.

Falk, J. L. (1971). The nature and determinants of adjunctive behavior. *Physiology and Behavior*, *6*, 577–588.

Falkenberg, C. A., & Barbetta, P. M. (2013). The effects of a self-monitoring package on homework completion and accuracy of students with disabilities in an inclusive general education classroom. *Journal of Behavioral Education*, *22*(3), 190–210. doi:10.1007/s10864-013-9169-1

Fantino, E., Preston, R. A., & Dunn, R. (1993). Delay reduction: Current status. *Journal of the Experimental Analysis of Behavior*, *60*, 159–169.

Fantuzzo, J. W., & Clement, P. W. (1981). Generalization of the effects of teacher and self-administered token reinforcers to nontreated students. *Journal of Applied Behavior Analysis*, *14*, 435–447.

Fantuzzo, J. W., Rohrbeck, C. A., Hightower, A. D., & Work, W. C. (1991). Teacher's use and children's preferences of rewards in elementary school. *Psychology in the Schools*, *28*, 175–181.

Farber, R. S., Dickson, C. A., & Dube, W. V. (2017). Reducing overselective stimulus control with differential observing responses. *Journal of Applied Behavior Analysis*, *50*, 87–105. doi:10.1002/jaba.363

Farrimond, S. J., & Leland, L. S., Jr. (2006). Increasing donations to supermarket food-bank bins using proximal prompts. *Journal of Applied Behavior Analysis*, *39*, 249–251.

Favell, J. E., & McGimsey, J. E. (1993). Defining an acceptable treatment environment. In R. Van Houten & S. Axelrod (Eds.), *Behavior analysis and treatment* (pp. 25–45). New York: Plenum Press.

Favell, J. E., Azrin, N. H., Baumeister, A. A., Carr, E. G., Dorsey, M. F., Forehand, R., & Solnick, J. V. (1982). The treatment of self-injurious behavior. *Behavior Therapy*, *13*, 529–554.

Favell, J. E., McGimsey, J. F., & Jones, M. L. (1980). Rapid eating in the retarded: Reduction by nonaversive procedures. *Behavior Modification*, *4*, 481–492.

Fawcett, S. B. (1991). Social validity: A note on methodology. *Journal of Applied Behavior Analysis*, *24*, 235–239.

Felce, D., & Emerson, E. (2000). Observational methods in assessment of quality of life. In T. Thompson, D. Felce, & F. J. Simmons (Eds.), *Behavioral observation: Technology and applications in developmental disabilities* (pp. 159–174). Baltimore: Paul H. Brookes.

Felixbrod, J. J., & O'Leary, K. D. (1973). Effects of reinforcement on children's academic behavior as a function of self-determined and externally imposed systems. *Journal of Applied Behavior Analysis*, *6*, 241–250.

Felixbrod, J. J., & O'Leary, K. D. (1974). Self-determination of academic standards by children: Toward freedom from external control. *Journal of Educational Psychology*, *66*, 845–850.

Ferrari, M., & Harris, S. (1981). The limits and motivational potential of sensory stimuli as reinforcers for autistic children. *Journal of Applied Behavior Analysis*, *14*, 339–343.

Ferreri, S. J., Allen, N., Hessler, T., Nobel, M., Musti-Rao, S., & Salmon, M. (2006). *Battling procrastination: Self-managing studying and writing for candidacy exams and dissertation defenses*. Symposium at 32nd Annual Convention of the Association for Behavior Analysis, Atlanta, GA.

Ferster, C. B. (1953). Sustained behavior under delayed reinforcement. *Journal of Experimental Psychology*, *45*, 218–224.

Ferster, C. B. (1958). Control of behavior in chimpanzees and pigeons by time out from positive reinforcement. *Psychological Monographs*, *72*(8), 1–38.

Ferster, C. B., & Perrott, M. C. (1968). *Behavior principles*. New York: Appleton-Century-Crofts.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. Englewood Cliffs, NJ: Prentice Hall.

Field, S. P., Frieder, J. E., Mcgee, H. M., Peterson, S. M., & Duinkerken, A. (2015). Assessing observer effects on the fidelity of implementation of functional analysis procedures. *Journal of Organizational Behavior Management*, *35*, 259–295. doi:10.1080/01608061.2015.1093058

Fields, L., Hobbie, S. A., Adams, B. J., & Reeve, K. F. (1999). Effects of training directionality and class size on equivalence class formation by adults. *The Psychological Record*, *49*, 703–724.

Fields, L., Travis, R., Roy, D., Yadlovker, E., Aguiar-Rocha, L., & Sturmey, P. (2009). Equivalence class formation: A method for teaching statistical interactions. *Journal of Applied Behavior Analysis*, *42*, 575–593.

Fienup, D. M., & Critchfield, T. S. (2010). Efficiently establishing concepts of inferential statistics and hypothesis decision making through contextually controlled equivalence classes. *Journal of Applied Behavior Analysis*, *43*, 437–462.

Fienup, D. M., & Critchfield, T. S. (2011). Transportability of equivalence-based programmed instruction: Efficacy and efficiency in a college classroom. *Journal of Applied Behavior Analysis*, *44*, 435–450.

Fienup, D. M., Covey, D. P., & Critchfield, T. S. (2010). Teaching brain-behavior relations economically with stimulus equivalence technology. *Journal of Applied Behavior Analysis*, *43*, 19–33.

Fienup, D. M., Critchfield, T. S., & Covey, D. P. (2009). Building contextually-controlled equivalence classes to teach about inferential statistics. A preliminary demonstration. *Experimental Analysis of Human Behavior Bulletin*, *27*, 1–10.

Fienup, D. M., Hamelin, J., Reyes-Giordano, K., & Falcomata, T. S. (2011). College-level instruction: Derived relations and programmed instruction. *Journal of Applied Behavior Analysis*, *44*, 413–416.

Finney, J. W., Putnam, D. E., & Boyd, C. M. (1998). Improving the accuracy of self-reports of adherence. *Journal of Applied Behavior Analysis*, *31*, 485–488.

Fiorile, C. A., & Greer, R. D. (2007). The induction of naming in children with no echoic-to-tact responses as a function of multiple exemplar instruction. *The Analysis of Verbal Behavior*, *23*, 71–88.

Fisch, G. S. (1998). Visual inspection of data revisited: Do the eyes still have it? *The Behavior Analyst*, *21*, 111–123.

Fischer, S., Iwata, B., & Worsdell, A. (1997). Attention as an establishing operation and as reinforcement during functional analyses. *Journal of Applied Behavior Analysis*, *30*, 335–338.

Fisher, R. A. (1925). *Statistical methods for research workers*. Edinburgh, Scotland: Oliver & Boyd.

Fisher, R. A. (1935). *The design of experiments*. London, UK: Oliver & Boyd.

Fisher, W., Piazza, C. C., Bowman, L. G., & Almari, A. (1996). Integrating caregiver report with a systematic choice assessment to enhance reinforcer identification. *American Journal on Mental Retardation*, *101*, 15–25.

Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis*, *25*, 491–498.

Fisher, W. W., & Iwata, B. A. (1996). On the function of self-restraint and its relationship to self-injury. *Journal of Applied Behavior Analysis*, *29*, 93–98.

Fisher, W. W., & Lerman, D. C. (2014). It has been said that, "There are three degrees of falsehoods: Lies, damn lies, and statistics." *Journal of School Psychology*, *52*, 135–140.

Fisher, W. W., Grace, N. C., & Murphy, C. (1996). Further analysis of the relationship between self-injury and self-restraint. *Journal of Applied Behavior Analysis*, *29*, 103–106.

Fisher, W. W., Greer, B. D., Romani, P. W., Zangrillo, A. N., & Owen, T. M. (2016). Comparisons of synthesized and individual reinforcement contingencies during functional analysis. *Journal of Applied Behavior Analysis*, *49*, 596–616. doi:10.1002/jaba.314

Fisher, W. W., Groff, R. A., & Roane, H. S. (2011). Applied Behavior Analysis: History, philosophy, principles, and basic methods. In W. W. Fisher, C. C. Piazza, and H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 3–16). New York: Guilford Press.

Fisher, W. W., Kelley, M. E., & Lomas, J. E. (2003). Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavior Analysis*, *36*, 387–406.

Fisher, W. W., Kuhn, D. E., & Thompson, R. H. (1998). Establishing discriminative control of responding using functional and alternative reinforcers during functional communication training. *Journal of Applied Behavior Analysis*, *31*, 543–560.

Fisher, W. W., Lindauer, S. E., Alterson, C. J., & Thompson, R. H. (1998). Assessment and treatment of destructive behavior maintained by stereotypic object manipulation. *Journal of Applied Behavior Analysis*, *31*, 513–527.

Fisher, W. W., Piazza, C. C., Bowman, L. G., Kurtz, P. F., Sherer, M. R. & Lachman, S. R. (1994). A preliminary evaluation of empirically derived consequences for the treatment of pica. *Journal of Applied Behavior Analysis*, *27*, 447–457.

Fisher, W. W., Piazza, C. C., Cataldo, M. E., Harrell, R., Jefferson, G., & Conner, R. (1993). Functional communication training with and without extinction and punishment. *Journal of Applied Behavior Analysis*, *26*, 23–36.

Fisher, W. W., Piazza, C. C., & Roane, H. S. (2011). *Handbook of applied behavior analysis*. New York: Guilford Press.

Fiske, K., & Delmolino, L. (2012). Use of discontinuous methods of data collection in behavioral intervention: Guidelines for practitioners. *Behavior Analysis in Practice*, *5*(2), 77–81.

Fitterling, J. M., Martin, J. E., Gramling, S., Cole, P., & Milan, M. A. (1988). Behavioral management of exercise training in vascular headache patients: An investigation of exercise adherence and headache activity. *Journal of Applied Behavior Analysis*, *21*, 9–19.

Flanagan, B., Goldiamond, I., & Azrin, N. (1958). Operant stuttering: The control of stuttering behavior through response-contingent consequences. *Journal of Experimental Analysis of Behavior*, *1*, 173–177.

Flaute, A. J., Peterson, S. M., Van Norman, R. K., Riffle, T., & Eakins, A. (2005). Motivate me! 20 tips for using a MotivAider® to improve your classroom. *Teaching Exceptional Children Plus*, 2(2), Article 3. Retrieved March 1, 2006, from http://escholarship.bc.edu/education/tecplus/vol2/iss2/art3

Fleece, L., Gross, A., O'Brien, T., Kistner, J., Rothblum, E., & Drabman, R. (1981). Elevation of voice volume in young developmentally delayed children via an operant shaping procedure. *Journal of Applied Behavior Analysis*, *14*, 351–355.

Flood, W. A., & Wilder, D. A. (2002). Antecedent assessment and assessment-based treatment of off-task behavior in a child diagnosed with attention deficit-hyperactivity disorder (ADHD). *Education and Treatment of Children*, *25*(3), 331–338.

Fogel, V. A., Miltenberger, R. G., Graves, R., & Koehler, S. (2010). Evaluating the effects of exergaming on physical activity among inactive children in a physical education classroom. *Journal of Applied Behavior Analysis*, *43*, 591–600.

Follette, W. C., & Darrow, S. M. (2014). Clinical behavior analysis. In F. K. McSweeney & E. S. Murphy (Eds.), *The Wiley-Blackwell handbook of operant and classical conditioning* (pp. 669–693). Oxford, UK: John Wiley & Sons.

Fonger, A. M., & Malott, R. W. (2018). Using shaping to teach eye contact to children with Autism Spectrum Disorder. *Behavior Analysis in Practice*, https://doi.org/10.1007/s40617-018-0245-9

Foster, S. L., & Mash, E. J. (1999). Assessing social validity in clinical treatment research issues and procedures. *Journal of Consulting and Clinical Psychology*, *67*, 308–319.

Foster, W. S. (1978). Adjunctive behavior: An underreported phenomenon in applied behavior analysis? *Journal of Applied Behavioral Analysis*, *11*, 545–546.

Fournier, A., Ehrhart, I. J., Glindemann, K. E., & Geller, E. S. (2004). Intervening to decrease alcohol abuse at university parties: Differential reinforcement of intoxication level. *Behavior Modification*, *28*, 167–181.

Fournier, A. K., & Berry, T. D. (2013). Interventions to increase hand hygiene behavior among university students: Implications for response cost and socially-assisted interventions. *Behavior and Social Issues*, *21*, 152–164.

Fowler, S. A., & Baer, D. M. (1981). "Do I have to be good all day?" The timing of delayed reinforcement as a factor in generalization. *Journal of Applied Behavior Analysis*, *14*, 13–24.

Fox, A. E., & Belding, D. L. (2015). Reducing pawing in horses using positive reinforcement. *Journal of Applied Behavior Analysis*, *48*, 936–940. doi:v10.1002/jaba.241

Fox, A. E., Bailey, S. R., Hall, E. G., & St. Peter, C. C. (2012). Reduction of biting and chewing of horses using differential reinforcement of other behavior. *Behavioural Processes*, *91*, 125–128. doi:v10.1016/j.beproc.2012.05.001

Fox, D. K., Hopkins, B. L., & Anger, W. K. (1987). The long-term effects of a token economy on safety performance in open-pit mining. *Journal of Applied Behavior Analysis*, *20*, 215–224.

Foxx, R. M. (1982). *Decreasing behaviors of severely retarded and autistic persons*. Champaign, IL: Research Press.

Foxx, R. M. (1996). Twenty years of applied behavior analysis in treating the most severe problem behavior: Lessons learned. *The Behavior Analyst*, *19*, 225–235.

Foxx, R. M., & Azrin, N. H. (1972). Restitution: A method of eliminating aggressive-disruptive behavior of retarded and brain damaged patients. *Behavior Research and Therapy*, *10*, 15–27.

Foxx, R. M., & Azrin, N. H. (1973). The elimination of autistic self-stimulatory behavior by overcorrection. *Journal of Applied Behavior Analysis*, *6*, 1–14.

Foxx, R. M., & Bechtel, D. R. (1983). Overcorrection: A review and analysis. In S. Axelrod & J. Apsche (Eds.), *The effects of punishment on human behavior* (pp. 133–220). New York: Academic Press.

Foxx, R. M., & Brown, R. A. (1979). Nicotine fading and self-monitoring for cigarette abstinence or controlled smoking. *Journal of Applied Behavior Analysis*, *12*, 111–115.

Foxx, R. M., & Mulick, J. A. (Eds.). (2016). *Controversial therapies for autism and intellectual disabilities: Fad, fashion, and science in professional practice* (2nd ed.). New York: Routledge/Taylor & Francis.

Foxx, R. M., & Rubinoff, A. (1979). Behavioral treatment of caffeinism: Reducing excessive coffee drinking. *Journal of Applied Behavior Analysis*, *12*, 335–344.

Foxx, R. M., & Shapiro, S. T. (1978). The timeout ribbon: A non-exclusionary timeout procedure. *Journal of Applied Behavior Analysis*, *11*, 125–143.

Fragale, C. L., O'Reilly, M. F., Aguilar, J., Pierce, N., Lang, R., Sigafoos, J., & Lancioni, G. (2012). The influence of motivating operations on generalization probes of specific mands by children with autism. *Journal of Applied Behavior Analysis*, *45*, 565–577.

Frampton, S. E., Wymer, S. C., Hansen, B., & Shillingsburg, M. A. (2016). The use of matrix training to promote generative language with children with autism. *Journal of Applied Behavior Analysis*, *49*, 869–883. doi.org/10.1002/jaba.340

France, K. G., & Hudson, S. M. (1990). Behavior management of infant sleep disturbance. *Journal of Applied Behavior Analysis*, *23*, 91–98.

Franklin, R. D., Gorman, B. S., Beasley, T. M., & Allison, D. B. (1996). Graphical display and visual analysis. In R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.), *Design and analysis of single-case research* (pp. 119–158). Hillsdale, NJ: Erlbaum.

Frederick, L. D., Davis, D. H., Alberto, P. A., & Waugh, R. E. (2013). From initial phonics to functional phonics: Teaching word-analysis skills to students with moderate intellectual disability. *Education and Training in Autism and Developmental Disabilities*, *48*, 49–66.

Freeland, J. T., & Noell, G. H. (1999). Maintaining accurate math responses in elementary school students: The effects of delayed intermittent reinforcement and programming common stimuli. *Journal of Applied Behavior Analysis*, *32*, 211–215.

Freeland, J. T., & Noell, G. H. (2002). Programming for maintenance: An investigation of delayed intermittent reinforcement and common stimuli to create indiscriminable contingencies. *Journal of Behavioral Education*, *11*, 5–18.

Freeman, J., Simonsen, B., McCoach, D. B., Sugai, G., Lombardi, A., & Horner, R. (2016). Relationship between school-wide positive behavior interventions and supports and academic, attendance, and behavior outcomes in high schools. *Journal of Positive Behavior Interventions*, *18*, 41–51.

Freeman, K. A. (2006). Treating bedtime resistance with the bedtime pass: A systematic replication and component analysis with 3-year-olds. *Journal of Applied Behavior Analysis*, *39*, 423–428.

Freeze Denton, T. & Meindl, J. N. (2016). The effect of colored overlays on reading fluency in individuals with dyslexia. *Behavior Analysis in Practice*, *9*, 191–198. doi 10.1007/s40617-015-0079-7

Friedling, C., & O'Leary, S. G. (1979). Effects of self-instructional training on second and third-grade hyperactive children: A failure to replicate. *Journal of Applied Behavior Analysis*, *12*, 211–219.

Friman, P. C. (1990). Nonaversive treatment of high-rate disruptions: Child and provider effects. *Exceptional Children*, *57*, 64–69.

Friman, P. C. (2000). "Transitional objects" as establishing operations for thumb sucking: A case study. *Journal of Applied Behavior Analysis*, *33*, 507–509.

Friman, P. C. (2006). The future of applied behavior analysis is under the dome. *ABA Newsletter*, *29*, 4–7.

Friman, P. C. (2017). You are in the way! Opening lines of transmission for Skinner's view of behavior. *The Behavior Analyst*, *40*, 173–177.

Friman, P. C., & Jones, K. M. (2005). Behavioral treatment for nocturnal enuresis. *Journal of Early and Intensive Behavioral Intervention*, *2*, 259–267.

Friman, P. C., & Piazza, C. C. (2011). Behavioral pediatrics: Integrating applied behavior analysis with pediatric medicine. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 433–450). New York: Guilford Press.

Friman, P. C., & Poling, A. (1995). Making life easier with effort: Basic findings and applied research on response effort. *Journal of Applied Behavior Analysis*, *28*, 538–590.

Friman, P. C., Finney, J. W., Rapoff, M. A., & Christophersen, E. R. (1985). *Journal of Applied Behavior Analysis*, *18*, 315–321.

Friman, P. C., Hayes, S. C., & Wilson, K. G. (1998). Why behavior analysts should study emotion: The example of anxiety. *Journal of Applied Behavior Analysis*, *31*, 137–156.

Friman, P. C., Hoff, K. E., Schnoes, C., Freeman, K. A., Woods, D. W., & Blum, N. (1999). The bedtime pass: An approach to bedtime crying and leaving the room. *Archives of Pediatric and Adolescent Medicine*, *153*, 1027–1029.

Fritz, J. N., Iwata, B. A., Rolider, N. U., Camp, E. M., & Neidert, P. L. (2012). Analysis of self-recording in self-management interventions for stereotypy. *Journal of Applied Behavior Analysis*, *45*(1), 55–68. http://doi.org/10.1901/jaba.2012.45-55

Frost, L. A., & Bondy, A. S. (1994). *The picture exchange communication system training manual*. Cherry Hill, NJ: Pyramid Educational Consultants, Inc.

Fryling, M. J., & Baires, N. A. (2016). The practical importance of the distinction between open and closed-ended indirect assessments. *Behavior Analysis in Practice*, *9*(2), 146–151. doi 10.1007/s40617-016-0115-2

Fryling, M. J., Wallace, M. D., & Yassine, J. N. (2012). Impact of treatment integrity on intervention effectiveness. *Journal of Applied Behavior Analysis*, *45*, 449–453.

Fu, S. B., Penrod, B., Fernand, J. K., Whelan, C. M., Griffith, K., & Medved, S. (2015). The effects of modeling contingencies in the treatment of food selectivity in children with autism. *Behavior Modification*, *39*(6), 771–784.

Fuhrman, A. M., Fisher, W. W., & Greer, B. D. (2016). A preliminary investigation on improving functional communication training by mitigating resurgence of destructive behavior. *Journal of Applied Behavior Analysis*, *49*, 884–899. doi:10.1002/jaba.338

Fuller, P. R. (1949). Operant conditioning of a vegetative organism. *American Journal of Psychology*, *62*, 587–590.

Fullerton, E. K., Conroy, M. A., & Correa, V. I. (2009). Early childhood teachers' use of specific praise statements with young children at risk for behavioral disorders. *Behavioral Disorders*, *34*, 118–135.

Furlong, M. J., & Wampold, B. E. (1982). Intervention effects and relative variation as dimensions in experts' use of visual inference. *Journal of Applied Behavior Analysis*, *15*, 415–421.

Furman, T. M., & Lepper, T. L. (2018). Applied behavior analysis: Definitional difficulties. *The Psychological Record*, *68*, 103–105. doi.org/10.1007/s40732-018-0282-3

Gable, R. A., Arllen, N. L., & Hendrickson, J. M. (1994). Use of students with emotional/behavioral disorders as behavior change agents. *Education and Treatment of Children*, *17*(3), 267–276.

Gabor, A. M. Fritz, J. N., Roath, C. T., Rothe, B. R., & Gourley, D. A. (2016). Caregiver preference for reinforcement-based interventions for problem behavior maintained by positive reinforcement. *Journal of Applied Behavior Analysis*, *49*, 6215–6227. doi:10.1002/jaba.286

Galbicka, G. (1994). Shaping in the 21st century: Moving percentile schedules into applied settings. *Journal of Applied Behavior Analysis*, *27*, 739–760.

Galbicka, G., & Platt, J. R. (1984). Interresponse-time punishment: A basis for shock-maintained behavior. *Journal of the Experimental Analysis of Behavior*, *41*, 291–308.

Galizio, M., & Bruce, K. E. (2017). Abstraction, multiple exemplar training and the search for derived stimulus relations in animals. *Perspectives on Behavioral Science*, https://doi.org/10.1007/s40614-017-0112-y

Galizio, M., & Bruce, K. E. (2018). Abstraction, multiple exemplar training, and the search for derived stimulus relations in animals. *Perspectives on Behavior Science*, *41*, 45–67.

Gallager, S. M., & Keenan, M. (2015). Stimulus equivalence and the Mini Mental Status Examination in the elderly. *European Journal of Behavior Analysis*, *10*, 159–165.

Gamba, J., Goyos, C., & Petursdottir, A. I. (2015). The functional independence of mands and tacts: Has it been demonstrated empirically? *The Analysis of Verbal Behavior*, *31*, 39–58. doi.org/10.1007/s40616-014-0026-7

Gambrill, E. (2003). Science and its use and neglect in the human services. In K. S. Budd & T. Stokes (Eds.), *A small matter of proof: The legacy of Donald M. Baer* (pp. 63–76). Reno, NV: Context Press.

Garcia, D., Dukes, C., Brady, M. P., Scott, J., & Wilson, C. L. (2016). Using modeling and rehearsal to teach fire safety to children with autism. *Journal of Applied Behavior Analysis*, *49*, 699–704. doi:10.1002/jaba.331

Garcia, E. E. (1976). The development and generalization of delayed imitation. *Journal of Applied Behavior Analysis*, *9*, 499.

Gardner, A. W., Spencer, T. D., Boelter, E. W., DuBard, M., & Jennett, H. K. (2012). A systematic review of brief functional analysis methodology with typically developing children. *Education and Treatment of Children*, *35*(2), 313–332.

Gardner, A. W., Wacker, D. P., & Boelter, E. W. (2009). An evaluation of the interaction between quality of attention and negative reinforcement with children who display escape-maintained problem behavior. *Journal of Applied Behavior Analysis*, *42*, 343–348.

Gardner, R., III, Heward, W. L., & Grossi, T. A. (1994). Effects of response cards on student participation and academic achievement: A systematic replication with inner-city students during whole-class science instruction. *Journal of Applied Behavior Analysis*, *27*, 63–71.

Garfinkle, A. N., & Schwartz, I. S. (2002). Peer imitation: Increasing social interactions in children with autism and other developmental disabilities in inclusive preschool classrooms. *Topics in Early Childhood Special Education*, *22*, 26–38.

Garner, K. (2012). Case study: The conscientious kid. In R. W. Malott, *I'll stop procrastinating when I get around to it: Plus other cool ways to succeed in school and life using behavior analysis to get your act together* (pp. 3–36, 3–37). Kalamazoo, MI: Behaviordelia.

Gast, D. L. (2014). General factors in measurement and evaluation. In D. L. Gast & J. R. Ledford (Eds.), *Single case research methodology applications in special education and behavioral sciences*. New York: Routledge.

Gast, D. L., & Ledford, J. R. (Eds.). (2014). *Single case research methodology: Applications in special education and behavioral sciences* (2nd ed.). New York: Routledge.

Gast, D. L., & Wolery, M. (1987). Severe maladaptive behaviors. In M. E. Snell (Ed.), *Systematic instruction of people with severe handicaps* (3rd ed.). Columbus, OH: Charles E. Merrill.

Gast, D. L., Jacobs, H. A., Logan, K. R., Murray, A. S., Holloway, A., & Long, L. (2000). Pre-session assessment of preferences for students with profound multiple disabilities. *Education and Training in Mental Retardation and Developmental Disabilities*, *35*, 393–405.

Gast, D. L., VanBiervliet, A., & Spradlin, J. E. (1979). Teaching number-word equivalences: A study of transfer. *American Journal of Mental Deficiency*, *83*, 524–527.

Gaylord-Ross, R. (1980). A decision model for the treatment of aberrant behavior in applied settings. In W. Sailer, B. Wilcox, & L. Brown (Eds.), *Methods of instruction for severely handicapped students* (pp. 135–158). Baltimore: Paul H. Brookes.

Gee, K., Graham, N., Goetz, L., Oshima, G., & Yoshioka, K. (1991). Teaching students to request the continuation of routine activities by using time delay and decreasing physical assistance in the context of chain interruption. *Journal of the Association for Persons with Severe Handicaps*, *10*, 154–167.

Geiger, K. B., Carr, J. E., & LeBlanc, L. A. (2010). Function-based treatments for escape-maintained problem behavior: A treatment selection model for practicing behavior analysts. *Behavior Analysis in Practice*, *3*, 22–32.

Geller, E. S. (2012). *Actively caring for people: Cultivating a culture of compassion*. Newport News, VA: Make-A-Difference, LLC.

Geller, E. S., Paterson, L., & Talbott, E. (1982). A behavioral analysis of incentive prompts for motivating seat belt use. *Journal of Applied Behavior Analysis*, *15*, 403–415.

Gena, A., Krantz, P. J., McClannahan, L. E., & Poulson, C. L. (1996). Training and generalization of affective behavior displayed by youth with autism. *Journal of Applied Behavior Analysis*, *29*, 291–304.

Gentile, J. R., Roden, A. H., & Klein, R. D. (1972). An analysis-of-variance model for the intrasubject replication design. *Journal of Applied Behavior Analysis*, *5*, 193–198.

Gewirtz, J. L., & Pelaez-Nogueras, M. (2000). Infant emotions under the positive-reinforcer control of caregiver attention and touch. In J. C. Leslie & D. Blackman (Eds.), *Issues in experimental and applied analyses of human behavior* (pp. 271–291). Reno, NV: Context Press.

Ghaemmaghami, M., Hanley, G. P., & Jessel, J. (2016). Contingencies promote delay tolerance. *Journal of Applied Behavior Analysis*, *49*, 548–575. doi:10.1002/jaba.333

Gil-Luciano, B., Ruiz, F. J., Valdivia-Salas, S., & Suárez-Falcón, J. C. (2017). Effect of framing behavior through deictic/hierarchical relations and specifying augmental functions in promoting psychological flexibility. *The Psychological Record*, *67*, 1–9.

Gil, K. M., Collins, F. L., & Odom, J. V. (1986). The effects of behavioral vision training on multiple aspects of visual functioning in myopic adults. *Journal of Behavioral Medicine*, *9*(4), 373–387.

Gilic, L., & Greer, R. D. (2011). Establishing naming in typically developing two-year-old children as a function of multiple exemplar speaker and listener experiences. *The Analysis of Verbal Behavior*, *27*, 157–177. doi.org/10.1007/BF03393099

Glenn, S. S. (2004). Individual behavior, culture, and social change. *The Behavior Analyst*, *27*, 133–151.

Glenn, S. S., Ellis, J., & Greenspoon, J. (1992). On the revolutionary nature of the operant as a unit of behavioral selection. *American Psychologist*, *47*, 1329–1336.

Glynn, E. L. (1970). Classroom applications of self-determined reinforcement. *Journal of Applied Behavior Analysis*, *3*, 123–132.

Glynn, E. L., Thomas, J. D., & Shee, S. M. (1973). Behavioral self-control of on-task behavior in an elementary classroom. *Journal of Applied Behavior Analysis*, *6*, 105–114.

Glynn, S. M. (1990). Token economy: Approaches for psychiatric patients: Progress and pitfalls over 25 years. *Behavior Modification*, *14*(4), 383–407.

Goddard, Y. L., & Heron, T. E. (1998, July/Aug). Please teacher, help me learn to spell better—Teach me self-correction. *Teaching Exceptional Children*, *30*(6), 38–43.

Goetz, E. M., & Baer, D. M. (1973). Social control of form diversity and the emergence of new forms in children's blockbuilding. *Journal of Applied Behavior Analysis*, *6*, 209–217.

Goetz, L., Gee, K., & Sailer, W. (1985). Using a behavior chain interruption strategy to teach communication skills to students with severe disabilities. *Research and Practice for Persons with Severe Disabilities*, *10*, 21–30.

Goetz, L., Gee, K., & Sailer, W. (1985). Using a behavior chain interruption strategy to teach communication skills to students with severe disabilities. *Journal of the Association for Persons with Severe Handicaps*, *10*, 21–30.

Goh, H., Iwata, B. A., & Kahng, S. (1999). Multicomponent assessment and treatment of cigarette pica. *Journal of Applied Behavior Analysis*, *32*, 297–316. doi.org/10.1901/jaba.1999.32–297

Goh, H.-L., & Iwata, B. A. (1994). Behavioral persistence and variability during extinction of self-injury maintained by escape. *Journal of Applied Behavior Analysis*, *27*, 173–174.

Goldiamond, I. (1965). Self-control procedures in personal behavior problems. *Psychological Reports*, *17*, 851–868.

Goldiamond, I. (1974). Toward a constructional approach to social problems: Ethical and constitutional issues raised by applied behavior analysis. *Behaviorism*, *2*, 1–85.

Goldiamond, I. (1976a). Self-reinforcement. *Journal of Applied Behavior Analysis*, *9*, 509–514.

Goldiamond, I. (1976b). Fables, armadyllics, and self-reinforcement. *Journal of Applied Behavior Analysis*, *9*, 521–525.

Goldstein, H. (1983). Recombinative generalization: Relationships between environmental conditions and the linguistic repertoires of language learners. *Analysis and Intervention in Developmental Disabilities*, *3*, 279–293. doi:10.1016/0270-4684(83)90002-2

Goldwater, B. C., & Acker, L. E. (1995). A descriptive notation system for contingency diagramming in behavior analysis. *The Behavior Analyst*, *18*, 113–121.

Goodall, J. (1991) *Through a window: My thirty years with the chimpanzees* of *Gombe*. New York: Houghton Mifflin.

Goodall, K. (1972, November). Shapers at work. *Psychology Today*, *5*, 53–63, 132–138.

Gorn, G. J. (1982). The effects of music in advertising on choice behavior: A classical conditioning approach. *Journal of Marketing*, *46*, 94–101.

Gottlieb, D. A., & Begej, E. L. (2014). Principles of Pavlovian conditioning: Description, content, function. In F. K. McSweeney & E. S. Murphy (Eds.), *The Wiley-Blackwell handbook of operant and classical conditioning* (pp. 3–25). West Sussex, UK: John Wiley & Sons.

Gottschalk, J. M., Libby, M. E., & Graff, R. B. (2000). The effects of establishing operations on preference assessment outcomes. *Journal of Applied Behavior Analysis*, *33*, 85–88.

Goyos, A. C. N., Souza, S. R., Silvares, E. F. M., & Saunders, R. R. (2007). Emergences of printing and spelling skills from constructed-response matching-to-sample instruction (CRMTS). *European Journal of Behavior Analysis*, *8*, 49–64.

Grace, N. C., Kahng, S., & Fisher, W. W. (1994). Balancing social acceptability with treatment effectiveness of an intrusive procedure: A case report. *Journal of Applied Behavior Analysis*, *27*, 171–172.

Grace, R. C., & Savastano, H. I. (2000). Temporal context and conditioned reinforcement value. *Journal of Experimental Psychology: General*, *129*, 427.

Graf, S. A., & Lindsley, O. R. (2002). *Standard Celeration Charting 2002*. Poland, OH: Graf Implements.

Graff, R. B., & Karsten, A. M. (2012a). Assessing preferences of individuals with developmental disabilities: A survey of current practices. *Behavior Analysis in Practice*, *5*(2), 37–48.

Graff, R. B., & Karsten, A. M. (2012b). Evaluation of a self-instruction package for conducting stimulus preference assessments. *Journal of Applied Behavior Analysis*, *45*, 69–82. doi:org/10.1901/jaba.2012.45-69

Grannan, L., & Rehfeldt, R. A. (2012). Emergent intraverbal responses via tact and match-to-sample instruction. *Journal of Applied Behavior Analysis*, *45*, 601–605.

Gravlee, C. C., Zenk, S. N., Woods, S., Rowe, Z., & Schultz, A. J. (2006). Handheld computers for direct observation of the social and physical environment. *Field Methods 18*, 382–397.

Gray, J. A. (1979). *Ivan Pavlov*. New York: Penguin Books.

Green, C. W., & Reid, D. H. (1996). Defining, validating, and increasing indices of happiness among people with profound multiple disabilities. *Journal of Applied Behavior Analysis*, *29*, 67–78.

Green, G., & Johnston, J. M. (2009). A primer on professional credentialing: Introduction to invited commentaries on licensing behavior analysts. *Behavior Analysis in Practice*, *2*(1), 51–52.

Green, G., & Saunders, R. (1998). Stimulus equivalence. In K. A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 229–262). New York: Plenum.

Green, G., Sigurdardottir, Z. G., & Saunders, R. R. (1991). The role of instructions in the transfer of ordinal functions through equivalence classes. *Journal of the Experimental Analysis of Behavior*, *55*, 287–304.

Green, L., & Estle, S. J. (2003). Preference reversals with food and water reinforcers in rats. *Journal of the Experimental Analysis of Behavior*, *79*(2), 233–242. doi:10.1901/jeab.2003.79-233

Green, L., & Freed, D. W. (1993). The substitutability of reinforcers. *Journal of the Experimental Analysis of Behavior*, *60*, 141–158.

Greene, B. F., Bailey, J. S., & Barber, F. (1981). An analysis and reduction of disruptive behavior on school buses. *Journal of Applied Behavior Analysis*, *14*, 177–192.

Greene, B. F., & Neistat, M. D. (1983). Behavior analysis in consumer affairs: Encouraging dental professionals to provide consumers with shielding from unnecessary X-ray exposure. *Journal of Applied Behavior Analysis*, *16*, 13–27.

Greenspan, S., & Negron, E. (1994). Ethical obligations of special services personnel. *Special Services in the Schools*, *8*(2), 185–209.

Greenwood, C. R., & Maheady, L. (1997). Measurable change in student performance: Forgotten standard in teacher preparation? *Teacher Education and Special Education*, *20*, 265–275.

Greenwood, S., Perrin, A., & Duggan, M. (2016). Pew Research Center. *Social Media Update 2016*. [2016-11-11]. http://www.pewinternet.org/2016/11/11/social-media-update-2016/

Greer, B. D., Fisher, W. W., Romani, P. W., & Saini, V. (2016). Behavioral momentum theory: A tutorial on response persistence. *The Behavior Analyst*, *39*, 269–291. doi:10.1007/s40614-016-0050-0

Greer, B. D., Fisher, W. W., Saini, V., Owen, T. M., & Jones, J. K. (2016). Functional communication training during reinforcement schedule thinning: An analysis of 25 applications. *Journal of Applied Behavior Analysis*, *49*, 105–121. doi:10.1002/jaba.265

Greer, B. D., Neibert, P. L., & Dozier, C. L. (2016). A component analysis of toilet-training procedures recommended for young children. *Journal of Applied Behavior Analysis*, *49*, 69–84. doi:10.1002/jaba.275

Greer, B. D., Neidert, P. L., Dozier, C. L., Payne, S. W., Zonneveld, K. L. M., & Harper, A. M. (2013). Functional analysis and treatment of problem behavior in early education classrooms. *Journal of Applied Behavior Analysis*, *46*, 289–295.

Greer, R. D. (1983). Contingencies of the science and technology of teaching and pre-behavioristic research practices in education. *Educational Researcher*, *12*, 3–9.

Greer, R. D., & Keohane, D. D. (2005). The evolution of verbal behavior in young children.

Greer, R. D., & Longano, J. (2010). A rose by naming: How we may learn how to do it. *The Analysis of Verbal Behavior*, *26*, 73–106.

Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis*. New York: Allyn & Bacon.

Greer, R. D., Dudek-Singer, J., & Gautreaux, G. (2006). Observational learning. *International Journal of Psychology*, *41*, 486–499. doi:10.1080/00207590500492435

Greer, R. D., Stolfi, L., Chavez-Brown, M., & Rivera-Valdes, C. (2005). The emergence of the listener to speaker component of naming in children as a function of multiple exemplar instruction. *The Analysis of Verbal Behavior*, *21*(1), 123–134.

Gresham, F. M. (1983). Use of a home-based dependent group contingency system in controlling destructive behavior: A case study. *School Psychology Review*, *12*(2), 195–199.

Gresham, F. M., Gansle, K. A., & Noell, G. H. (1993). Treatment integrity in applied behavior analysis with children. *Journal of Applied Behavior Analysis*, *26*, 257–263.

Grey, I., Healy, O., Leader, G., & Hayes, D. (2009). Using a Time Timer to increase appropriate waiting behavior in a child with developmental disabilities. *Research in Developmental Disabilities*, *30*, 359–366.

Griffee, K., & Dougher, M. J. (2002). Contextual control of stimulus generalization and stimulus equivalence in hierarchical categorization. *Journal of the Experimental Analysis of Behavior*, *78*, 433–447.

Griffen, A. K., Wolery, M., & Schuster J. W. (1992). Triadic instruction of chained food preparation responses: Acquisition and observational learning. *Journal of Applied Behavior Analysis*, *25*, 193–204.

Griffith, K. R., Ramos, A. L., Hill, K. E., & Miguel, C. F. (2018). Using equivalence-based instruction to teach piano skills to college students. *Journal of Applied Behavior Analysis*, *51*, 207–219. doi:10.1002/jaba.438

Griffith, R. G. (1983). The administrative issues: An ethical and legal perspective. In S. Axelrod & J. Apsche (Eds.), *The effects of punishment on human behavior* (pp. 317–338). New York: Academic Press.

Groskreutz, N. C., Groskreutz, M. P., Bloom, S. E., & Slocum, T. A. (2014). Generalization of negatively reinforced mands in children with autism. *Journal of Applied Behavior Analysis*, 47, 560–579.

Groskreutz, N. C., Groskreutz, M. P., & Higbee, T. S. (2011). Effects of varied levels of treatment integrity on appropriate toy manipulation in children with autism. *Research in Autism Spectrum Disorders*, *5*, 1358–1369.

Groskreutz, N. C., Karsina, A., Miguel, C. F., & Groskreutz, M. P. (2010). Using complex auditory-visual samples to produce emergent relations in children with autism. *Journal of Applied Behavior Analysis*, *43*, 131–136.

Gross, A. C., Fuqua, R. W., & Merritt, T. A. (2013). Evaluation of verbal behavior in older adults. *The Analysis of Verbal Behavior*, *29*, 85–99. doi.org/10.1007/BF03393126

Grossi, T. A. (1998). Using a self-operated auditory prompting system to improve the work performance of two employees with severe disabilities. *Journal of the Association for Persons with Severe Handicaps*, *23*, 149–154.

Grossi, T. A., & Heward, W. L. (1998). Using self-evaluation to improve the work productivity of trainees in a community-based restaurant training program. *Education and Training in Mental Retardation and Developmental Disabilities*, *33*, 248–263.

Grossi, T. A., Kimball, J. W., & Heward, W. L. (1994). What did you say? Using review of tape-recorded interactions to increase social acknowledgments by trainees in a community-based vocational program. *Research in Developmental Disabilities*, *15*, 457–472.

Grow, L., & LeBlanc, L. (2013). Teaching receptive language skills: Recommendations for instructors. *Behavior Analysis in Practice*, *6*(1), 56–75.

Grow, L. L., LeBlanc, L. A., & Carr, J. E. (2010). Developing stimulus control of the high-rate social approach responses of an adult with mental retardation: A multiple-schedule evaluation. *Journal of Applied Behavior Analysis*, *43*, 285–289. doi:10.1901/jaba.2010.43-285

Grunsell, J., & Carter, M. (2002). The behavior change interruption strategy: Generalization to out-of-routine contexts. *Education and Training in Mental Retardation and Developmental Disabilities*, *37*(4), 378–390.

Guercio, J. M., & Cormier, R. J. (2015). Blending fading procedures with forward chaining to address treatment resistance in an adult with an autism spectrum disorder. *Behavior Analysis in Practice*, *8*, 215–218.

Guercio, J. M. (2018). The importance of a deeper knowledge of the history and theoretical foundations of behavior analysis: 1863–1960. *Behavior Analysis: Research and Practice*, *18*, 4–15. doi.org/10.1037/bar0000123

Guercio, J. M., Podolska-Schroeder, H., & Rehfeldt, R. A. (2003). Using stimulus equivalence technology to teach emotion recognition to adults with acquired brain injury. *Brain Injury*, *18*, 593–601.

Guess, D., Sailer, W., Rutherford, G. & Baer, D. M. (1968). An experimental analysis of linguistic development: The productive use of the plural morpheme. *Journal of Applied Behavior Analysis*, *1*, 297–306.

Guilford, J. P. (1965). *Fundamental statistics in psychology and education*. New York: McGraw-Hill.

Guinther, P. M. (2017). Contextual influence over deriving others' true beliefs using a relational triangulation perspective-taking protocol (RT-PTP-M1). *Journal of the Experimental Analysis of Behavior*, *108*, 433–456.

Guinther, P. M., & Dougher, M. J. (2010). Semantic false memories in the form of derived relational intrusions following training. *Journal of the Experimental Analysis of Behavior*, *93*, 329–347.

Gulchak, D. J. (2008). Using a mobile handheld computer to teach a student with an emotional and behavioral disorder to self-monitor attention. *Education and Treatment of Children*, *31*(4), 567–581.

Gunby, K. V., Carr, J. E., & LeBlanc, L. A. (2010). Teaching abduction-prevention skills to children with autism. *Journal of Applied Behavior Analysis*, *43*, 107–112. doi:10.1901/jaba.2010.43-107

Gunter, P. L., Venn, M. L., Patrick, J., Miller, K. A., & Kelly, L. (2003). Efficacy of using momentary time samples to determine on-task behavior of students with emotional/behavioral disorders. *Education and Treatment of Children*, *26*, 400–412.

Gurrad, A. M., Weber, K. P., & McLaughlin, T. F. (2002). The effects of contingency contracting for a middle school student with attention deficit hyperactivity disorder during corrective reading lessons: A case report. *International Journal of Special Education*, *17*, 26–31.

Gutierrez, A., Jr., Vollmer, T. R., Dozier, C. L., Borrero, J. C., Rapp, J. T., Bourret, J. C., & Gadaire, D. (2007). Manipulating establishing operations to verify and establish stimulus control during mand training. *Journal of Applied Behavior Analysis*, *40*, 645–658.

Gutowski, S. J., & Stromer, R. (2003). Delayed matching to two-picture samples by individuals with and without disabilities: An analysis of the role of naming. *Journal of Applied Behavior Analysis*, *36*, 487–505. doi:10.1901/jaba.2003.36-487

Guttman, N., & Kalish, H. (1956). Discriminability and generalization. *Journal of Experimental Psychology*, *51*, 79–88.

Haagbloom, S. J., Warnick, R., Warnick, J. E., Jones, V. K., Yarbrough, G. L., Russell, T. M., et al. (2002). The 100 most eminent psychologists of the 20th century. *Review of General Psychology*, *6*, 139–152.

Hackenberg, T. D. (1995). Jacques Loeb, B. F. Skinner, and the legacy of prediction and control. *The Behavior Analyst*, *18*, 225–236.

Hackenberg, T. D. (2009). Token reinforcement: A review and analysis. *Journal of the Experimental Analysis of Behavior*, *91*, 257–286.

Hackenberg, T. D. (2018). Token reinforcement: Translational research and application. *Journal of Applied Behavior Analysis*, *51*, 393–435. doi:10.1002/jaba.439

Haegele, K. M., McComas, J. J., Dixon, M., & Burns, M. K. (2011). Using a stimulus equivalence paradigm to teach numerals, English words, and Native American words to preschool-age children. *Journal of Behavioral Education*, *20*, 283–296.

Hager, K. D. (2012). Self-monitoring as a strategy to increase student teachers' use of effective teaching practices. *Rural Special Education Quarterly*, *31*(4), 9–17.

Hagopian, L. P., & Adelinis, J. D. (2001). Response blocking with and without redirection for the treatment of pica. *Journal of Applied Behavior Analysis*, *34*, 527–530.

Hagopian, L. P., & Thompson, R. H. (1999). Reinforcement of compliance with respiratory treatment in a child with cystic fibrosis. *Journal of Applied Behavior Analysis*, *32*, 233–236.

Hagopian, L. P., Boelter, E. W., & Jarmolowicz, D. P. (2011). Reinforcement schedule thinning following functional communication training: Review and recommendations. *Behavior Analysis in Practice*, *4*(1), 4–16.

Hagopian, L. P., Farrell, D. A., & Amari, A. (1996). Treating total liquid refusal with backward chaining and fading. *Journal of Applied Behavior Analysis, 29*, 573–575.

Hagopian, L. P., Fisher, W. W., Sullivan, M. T., Acquisto, J., & Leblanc, L. A. (1998). Effectiveness of functional communication training with and without extinction and punishment: A summary of 21 inpatient cases. *Journal of Applied Behavior Analysis, 31*, 211–235.

Hagopian, L. P., Fisher, W. W., Thompson, R. H., Owen-DeSchryver, J., Iwata, B. A., & Wacker, D. P. (1997). Toward the development of structured criteria for interpretation of functional analysis data. *Journal of Applied Behavior Analysis, 30*, 313–326. doi: 10.1901/jaba.1997.30-313

Hagopian, L. P., González, M. L., Taylor Rivet, T., Triggs, M., & Clark, S. B. (2011). Response interruption and differential reinforcement of alternative behavior for the treatment of pica. *Behavioral Interventions, 26*(4), 309–325.

Hagopian, L. P., Rooker, G. W., & Yenokyan, G. (2018). Identifying predictive behavioral markers: A demonstration using automatically reinforced self-injurious behavior. *Journal of Applied Behavior Analysis, 51*, 443–465.

Hagopian, L. P., Rush, K. S., Lewin, A. B., & Long, E. S. (2001). Evaluating the predictive validity of a single stimulus engagement preference assessment. *Journal of Applied Behavior Analysis, 34*, 475–485.

Hake, D. F., & Azrin, N. H. (1965). Conditioned punishment. *Journal of the Experimental Analysis of Behavior, 6*, 297–298.

Hake, D. F., Azrin, N. H., & Oxford, R. (1967). The effects of punishment intensity on squirrel monkeys. *Journal of the Experimental Analysis of Behavior, 10*, 95–107.

Hall, G. A., & Sundberg, M. L. (1987). Teaching mands by manipulating conditioned establishing operations. *The Analysis of Verbal Behavior, 5*, 41–53.

Hall, J. E., & Lunt, I. (2005). Global mobility for psychologists: The role of psychology organizations in the United States, Canada, Europe, and other regions. *The American Psychologist, 60*, 712–726.

Hall, R. V., & Copeland, R. (1972). The responsive teaching model: A first step in shaping school personnel as behavior modification specialists. *Proceedings of the Third Banff International Conference on Behavior Modification* (pp. 125–150). Champaign, IL: Research Press.

Hall, R. V., & Fox, R. G. (1977). Changing-criterion designs: An alternative applied behavior analysis procedure. In B. C. Etzel, J. M. LeBlanc, & D. M. Baer (Eds.), *New developments in behavioral research: Theory, method, and application* (pp. 151–166). Hillsdale, NJ: Erlbaum.

Hall, R. V., Axelrod, S., Foundopoulos, M., Shellman, J., Campbell, R. A., & Cranston, S. S. (1971). The effective use of punishment to modify behavior in the classroom. *Educational Technology, 11*(4), 25.

Hall, R. V., Cristler, C., Cranston, S. S., & Tucker, B. (1970). Teachers and parents as researchers using multiple baseline designs. *Journal of Applied Behavior Analysis, 3*, 247–255.

Hall, R. V., Delquadri, J. C., & Harris, J. (1977, May). *Opportunity to respond: A new focus in the field of applied behavior analysis*. Paper presented at the Midwest Association for Behavior Analysis, Chicago, IL.

Hall, R. V., Fox, R., Williard, D., Goldsmith, L., Emerson, M., Owen, M., . . . & Davis, R. (1970). Modification of disrupting and talking-out behavior with the teacher as observer and experimenter. Paper presented at the American Educational Research Association Convention, Minneapolis, MN.

Hall, R. V., Lund, D., & Jackson, D. (1968). Effects of teacher attention on study behavior. *Journal of Applied Behavior Analysis, 1*, 1–12.

Hall, R. V., Panyan, M., Rabon, D., & Broden. D. (1968). Instructing beginning teachers in reinforcement procedures which improve classroom control. *Journal of Applied Behavior Analysis, 1*, 315–322.

Hall, S. S., DeBernardis, G. M., & Reiss, A. L. (2006). The acquisition of stimulus equivalence in individuals with fragile X syndrome. *Journal of Intellectual Disability Research, 50*, 643–651.

Hall, S. S., Maynes, N. P., & Reiss, A. L. (2009). Using percentile schedules to increase eye contact in children with fragile X syndrome. *Journal of Applied Behavior Analysis, 42*, 171–176.

Halldórsdóttir, L. Y., Valdimarsdóttir, H., & Arntzen, E. (2017). Effects of a multicomponent intervention to improve reading fluency in adult students. *Behavioral Interventions, 32*, 3–20. doi:10.1002/bin.1452

Hamblin, R. L., Hathaway, C., & Wodarski, J. S. (1971). Group contingencies, peer tutoring, and accelerating academic achievement. In E. A. Ramp & B. L. Hopkins (Eds.), *A new direction for education: Behavior analysis* (Vol. 1, pp. 41–53). Lawrence, KS: University of Kansas.

Hamlet, C. C., Axelrod, S., & Kuerschner, S. (1984). Eye contact as an antecedent to compliant behavior. *Journal of Applied Behavior Analysis, 17*, 553–557.

Hammer-Kehoe, J. (2012). Case study: Yoga. In R. W. Malott, *I'll stop procrastinating when I get around to it: Plus other cool ways to succeed in school and life using behavior analysis to get your act together* (pp. 6–74). Kalamazoo, MI: Behaviordelia.

Hammond, J. L., Iwata, B. A., Fritz, J. N., & Dempsey, C. M. (2011). Evaluation of fixed momentary DRO schedules under signaled and unsignaled arrangements. *Journal of Applied Behavior Analysis, 44*, 69–81. doi:10.1901/jaba.2011.44-69

Hanley, G. P. (2012). Functional assessment of problem behavior: Dispelling myths, overcoming implementation obstacles, and developing new lore. *Behavior Analysis in Practice, 5*(1), 54–72. http://doi.org.libproxy.library.wmich.edu/10.1007/BF03391818

Hanley, G. P. (2017). Editor's note. *Journal of Applied Behavior Analysis, 50*, 3–7.

Hanley, G. P., & Tiger, J. H. (2011). Differential reinforcement procedures. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 229–249). New York: Guilford Press.

Hanley, G. P., Cammilleri, A. P., Tiger, J. H., & Ingvarsson, E. T. (2007). A method for describing preschoolers' activity preferences. *Journal of Applied Behavior Analysis, 40*, 603–618.

Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*, 147–185. doi:10.1901/jaba.2003.36-147

Hanley, G. P., Iwata, B. A., & Roscoe, E. M. (2006). Some determinants of changes in preference over time. *Journal of Applied Behavior Analysis, 39*, 189–202. doi:10.1901/jaba.2006.163-04

Hanley, G. P., Iwata, B. A., & Thompson, R. H. (2001). Reinforcement schedule thinning following treatment with functional communication training. *Journal of Applied Behavior Analysis, 34*, 17–38.

Hanley, G. P., Iwata, B. A., Thompson, R. H., & Lindberg, J. S. (2001). A component analysis of "stereotypy and reinforcement" for alternative behavior. *Journal of Applied Behavior Analysis, 33*, 299–308.

Hanley, G. P., Jin, C. S., Vanselow, N. R., & Hanratty, L. A. (2014). Producing meaningful improvements in problem behavior of children with autism via synthesized analyses and treatments. *Journal of Applied Behavior Analysis, 47*, 16–36. doi:10.1002/jaba.106

Hanley, G. P., Piazza, C. C., Fisher, W. W., Contrucci, S. A., & Maglieri, K. A. (1997). Evaluation of client preference for function-based treatment packages. *Journal of Applied Behavior Analysis, 30*, 459–473. doi:10.1901/jaba.1997.30-459

Hanley, G. P., Piazza, C. C., Fisher, W. W., & Maglieri, K. A. (2005). On the effectiveness of and preference for punishment and extinction components of function-based interventions. *Journal of Applied Behavior Analysis, 38*, 51–65. doi:10.1901/jaba.2005.6-04

Hann, K. E., & McCracken, L. M. (2014). A systematic review of randomized controlled trials of Acceptance and Commitment Therapy for adults with chronic pain: Outcome domains, design quality, and efficacy. *Journal of Contextual Behavioral Science, 3*, 217–227.

Hanna, E. S., de Souza, D. G., de Rose, J. C., & Fonseca, M. (2004). Effects of delayed constructed-response identity matching on spelling of dictated words. *Journal of Applied Behavior Analysis, 37*, 223–227. doi:10.1901/jaba.2004.37-223

Hanna, E. S., Kohlsdorf, M., Quinteiro, R. S., de Melo, R. M., de Souza, D. G., de Rose, J. C., & McIlvane, W. J. (2011). Recombinative reading derived from pseudoword instruction in a miniature linguistic system. *Journal of the Experimental Analysis of Behavior, 95*, 21–40.

Hanratty, L., Miltenberger, R., & Florentino, S. (2016). Evaluating the effectiveness of a teaching package utilizing behavioral skills training and in situ training to teach gun safety skills in a preschool classroom. *Journal of Behavioral Education, 25*, 310–323.

Hansard, C., & Kazemi, E. (2018). Evaluation of video self-instruction for implementing paired-stimulus preference assessments. *Journal of Applied Behavior Analysis, 51*, 675–680.

Hansen, B. D., & Wills, H. P. (2014). The effects of goal setting, contingent reward, and instruction on writing skills. *Journal of Applied Behavior Analysis, 47*, 171–175. doi:10.1002/jaba.92

Harding, J., Wacker, D. P., Cooper, L. J., Millard, T., & Jensen-Kovalan, P. (1994). Brief hierarchical assessment of potential treatment components

with children in an outpatient clinic. *Journal of Applied Behavior Analysis*, *27*, 291–300. doi:10.1901/jaba.1994.27-291

Harding, J. W., Wacker, D. P., Berg, W. K., Rick, G., & Lee, J. F. (2004). Promoting response variability and stimulus generalization in martial arts training. *Journal of Applied Behavior Analysis*, *37*, 185–195.

Haring, T. G., & Kennedy, C. H. (1990). Contextual control of problem behavior. *Journal of Applied Behavior Analysis*, *23*, 235–243.

Harlow, H. F. (1949). The formation of learning sets. *Psychological Review*, *56*, 51–65.

Harlow, H. R. (1959). Learning set and error factor theory. In S. Kock (Ed.), *Psychology: A study of science* (Vol. 2, pp. 492–537). New York: McGraw-Hill.

Haroff, L. (2012). Case study: Exercise. In R. W. Malott, *I'll stop procrastinating when I get around to it: Plus other cool ways to succeed in school and life using behavior analysis to get your act together* (pp. 4–51). Kalamazoo, MI: Behaviordelia.

Harrell, J. P. (2012). Case study: Taking time to relax. In R. W. Malott, *I'll stop procrastinating when I get around to it: Plus other cool ways to succeed in school and life using behavior analysis to get your act together* (pp. 6–70). Kalamazoo, MI: Behaviordelia.

Harriage, B., Blair, K, & Miltenberger, R. (2016). An evaluation of a parent-implemented in situ pedestrian safety skills intervention for individuals with autism. *Journal of Autism and Developmental Disorders*, *46*, 2017–2027.

Harris, F. N., & Jenson, W. R. (1985). Comparisons of multiple-baseline across persons designs and AB designs with replication: Issues and confusions. *Behavioral Assessment*, *7*, 121–129.

Harris, F. R., Johnston, M. K., Kelly, C. S., & Wolf, M. M. (1964). Effects of positive social reinforcement on regressed crawling of a nursery school child. *Journal of Educational Psychology*, *55*, 35–41.

Harris, K. R., Friedlander, B. D., Saddler, B., Frizzelle, R., & Graham, S. (2005). Self-monitoring of attention versus self-monitoring of academic performance. *The Journal of Special Education*, *39*, 145–156.

Harris, R. (2008). *The happiness trap: How to stop struggling and start living*. Boston: Trumpeter.

Harris, S. L., & Wolchik, S. A. (1979). Suppression of self-stimulation: Three alternative strategies. *Journal of Applied Behavior Analysis*, *12*, 185–198.

Harrison, A. M., & Pyles, D. A. (2013). The effects of verbal instruction and shaping to improve tackling by high school football players. *Journal of Applied Behavior Analysis*, *46*, 518–522.

Hart B., & Risley T. R. (1975). Incidental teaching of language in the preschool. *Journal of Applied Behavior Analysis*, *8*, 411–420.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: Brookes.

Hart, B., Reynolds, N. J., Baer, D. M., Brawley, E. R., & Harris, F. R. (1968). Effect of contingent and non-contingent social reinforcement on the cooperative play of a preschool child. *Journal of Applied Behavior Analysis*, *1*, 73–76.

Hart, B. M., Allen, K. E., Buell, J. S., Harris, F. R., & Wolf, M. M. (1964). Effects of social reinforcement on operant crying. *Journal of Experimental Child Psychology*, *1*, 145–153.

Hartmann, D. P. (1974). Forcing square pegs into round holes: Some comments on "an analysis-of-variance model for the intrasubject replication design". *Journal of Applied Behavior Analysis*, *7*, 635–638.

Hartmann, D. P. (1977). Considerations in the choice of interobserver reliability estimates. *Journal of Applied Behavior Analysis*, *10*, 103–116.

Hartmann, D. P., & Hall, R. V. (1976). The changing criterion design. *Journal of Applied Behavior Analysis*, *9*, 527–532.

Harvey, M. T., May, M. E., & Kennedy, C. H. (2004). Nonconcurrent multiple baseline designs and the evaluation of educational systems. *Journal of Behavioral Education*, *13*, 267–276.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.

Hausman, N. L., Borrero, J. C., Fisher, A., & Kahng, S. W. (2014). Improving accuracy of portion-size estimations through a stimulus equivalence paradigm. *Journal of Applied Behavior Analysis*, *47*, 485–499.

Hawkins, R. P. (1979). The functions of assessment. *Journal of Applied Behavior Analysis*, *12*, 501–516.

Hawkins, R. P. (1984). What is "meaningful" behavior change in a severely/profoundly retarded learner: The view of a behavior analytic parent. In W. L. Heward, T. E. Heron, D. S. Hill, & J. Trap-Porter (Eds.), *Focus on behavior analysis in education* (pp. 282–286). Columbus, OH: Charles E. Merrill.

Hawkins, R. P. (1986). Selection of target behaviors. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 331–385). New York: Guilford Press.

Hawkins, R. P. (1991). Is social validity what we are interested in? *Journal of Applied Behavior Analysis*, *24*, 205–213.

Hawkins, R. P., & Anderson, C. M. (2002). On the distinction between science and practice: A reply to Thyer and Adkins. *The Behavior Analyst*, *26*, 115–119.

Hawkins, R. P., & Dobes, R. W. (1977). Behavioral definitions in applied behavior analysis: Explicit or implicit? In B. C. Etzel, J. M. LeBlanc, & D. M. Baer (Eds.), *New developments in behavioral research: Theory, method, and application* (pp. 167–188). Hillsdale, NJ: Lawrence Erlbaum.

Hawkins, R. P., & Dotson, V. A. (1975). Reliability scores that delude: An Alice in Wonderland trip through the misleading characteristics of interobserver agreement scores in interval recording. In E. Ramp & G. Semp (Eds.), *Behavior analysis: Areas of research and application* (pp. 359–376). Upper Saddle River, NJ: Prentice Hall.

Hawkins, R. P., Meadowcroft, P., Trout, B. A., & Luster, W. C. (1985). Foster family-based treatment. *Journal of Clinical Child Psychology*, *14*, 220–228.

Hayes, J., & Stewart, I. (2016). Comparing the effects of derived relational training and computer coding on intellectual potential in school-age children. *British Journal of Educational Psychology*, *86*, 397–411.

Hayes, L. B., & Van Camp, C. M. (2015). Increasing physical activity of children during school recess. *Journal of Applied Behavior Analysis*, *48*, 690–695. doi:10.1002/jaba.222

Hayes, L. J., Adams, M. A., & Rydeen, K. L. (1994). Ethics choice and value. In L. J. Hayes, G. Hayes, S. Moore, and P. Ghezzi (Eds.), *Ethical issues in developmental disabilities*. Reno, NV: Context Press.

Hayes, L. J., Thompson, S., & Hayes, S. C. (1989). Stimulus equivalence and rule following. *Journal of the Experimental Analysis of Behavior*, *52*, 275–291.

Hayes, L. J., Tilley, K. L., & Hayes, S. C. (1988). Extending equivalence class membership to gustatory stimuli. *The Psychological Record*, *38*, 473–482.

Hayes, S. C. (1984). Making sense of spirituality. *Behaviorism*, *12*, 99–110.

Hayes, S. C. (1985). Natural multiple baselines across persons: A reply to Harris and Jenson. *Behavioral Assessment*, *7*, 129–132.

Hayes, S. C. (Ed.). (1989). *Rule-governed behavior: Cognition, contingencies, and instructional control*. Reno, NV: Context Press.

Hayes, S. C. (2001). The greatest dangers facing behavior analysis today. *The Behavior Analyst Today*, *2*, 61–63.

Hayes, S. C. (2005). *Get out of your mind and into your life: The new acceptance and commitment therapy*. Oakland, CA: New Harbinger Publications.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York: Kluwer/Plenum.

Hayes, S. C., Fox, E., Gifford, E.V., Wilson, K.G., Barnes-Holmes, D., & Healy, O. (2001). Derived relational responding as learned behavior. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 21–49). New York: Plenum.

Hayes, S. C., Gifford, E. V., & Ruckstahl, L. E. (1998). Relational frame theory and executive function: A behavioral approach. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function* (pp. 279–305). Baltimore: Paul H. Brookes.

Hayes, S. C., Kohlenberg, B., & Hayes, L. J. (1991). The transfer of specific and general consequential functions through simple and conditional equivalence relations. *Journal of the Experimental Analysis of Behavior*, *56*, 119–137.

Hayes, S. C., Luoma, J., Bond, F., Masuda, A., & Lissis, J. (2007). Acceptance and commitment therapy: Model, processes, and outcomes. *Behaviour Research and Therapy*, *44*, 1–25.

Hayes, S. C., Rincover, A., & Solnick, J. V. (1980). The technical drift of applied behavior analysis. *Journal of Applied Behavior Analysis*, *13*, 275–285.

Hayes, S. C., Strosahl, K. D., & Wilson, K. W. (2016). *Acceptance and Commitment Therapy: The process and practice of mindful change* (2nd ed.). New York: Guilford Press.

Hayes, S. C., Zettle, R. D., & Rosenfarb, I. (1989). Rule-following. In S. C. Hayes (Ed.), *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 191–220). New York: Plenum.

Haynes, S. N., & Horn, W. F. (1982). Reactivity in behavioral observation: A methodological and conceptual critique. *Behavioral Assessment*, *4*, 369–385.

Hayward, E. O., & Homer, B. D. (2017). Reliability and validity of advanced theory-of-mind measures in middle childhood and adolescence. *British Journal of Developmental Psychology*, *35*, 454–462.

Heagle, A. I., & Rehfeldt, R. A. (2006). Teaching perspective-taking skills to typically developing children through derived relational responding. *Journal of Early and Intensive Behavioral Intervention*, *3*, 1–34.

Heal, N. A., & Hanley, G. P. (2007). Evaluating preschool children's preferences for motivational system during instruction. *Journal of Applied Behavior Analysis*, *40*, 249–261.

Heal, N. A., Hanley, G. P., & Layer, S. A. (2009). An evaluation of the relative efficacy of and children's preferences for teaching strategies that differ in amount of teacher directedness. *Journal of Applied Behavior Analysis*, *42*, 123–143. doi:10.1901/jaba.2009.42-123

Health Insurance Portability and Accountability Act of 1996 (HIPAA; Pub.L. 104-191, 110 Stat. 1936, enacted August 21, 1996).

Heckaman, K. A., Alber, S. R., Hooper, S., & Heward, W. L. (1998). A comparison of least-to-most prompts and progressive time delay on the disruptive behavior of students with autism. *Journal of Behavior Education*, *8*, 171–201.

Heffernan, L., & Lyons, D. (2016). Differential reinforcement of other behaviour for the reduction of severe nail biting. *Behavior Analysis in Practice*, *9*(3), 253–256.

Heinicke, M. R., Carr, J. E., Mozzoni, M. P., & Roane, S. (2009). Using differential reinforcement to decrease academic response latencies of an adolescent with acquired brain injury. *Journal of Applied Behavior Analysis*, *42*, 861–865.

Heldt, J., & Schlinger, H. D., Jr. (2012). Increased variability in tacting under a Lag 3 schedule of reinforcement. *The Analysis of Verbal Behavior*, *28*, 131–136.

Heron, T. E. (1978). Punishment: A review of the literature with implications for the teacher of mainstreamed children. *The Journal of Special Education*, *12*, 243–252.

Heron, T. E. (2014). *Instrument flying: 10 indispensible principles to know and remember*. Minneapolis, MN: Two Harbors Press.

Heron, T. E., & Heward, W. L. (1988). Ecological assessment: Implications for teachers of learning disabled students. *Learning Disability Quarterly*, *11*, 224–232.

Heron, T. E., Heward, W. L., Cooke, N. L., & Hill, D. S. (1983). Evaluation of a classwide peer tutoring system: First graders teach each other sight words. *Education and Treatment of Children*, *6*, 137–152.

Heron, T. E., Hippler, B. J., & Tincani, M. J. (2003). *How to help students complete classwork and homework assignments*. Austin, TX: PRO-ED.

Heron, T. E., Tincani, M. J., Peterson, S. M., & Miller, A. D. (2005). Plato's allegory of the cave revisited. Disciples of the light appeal to the pied pipers and prisoners in the darkness. In W. L. Heward, T. E. Heron, N. A. Neef, S. M. Peterson, D. M. Sainato, G. Cartledge, R. Gardner, III, L. D. Peterson, S. B. Hersh, & J. C. Dardig (Eds.), *Focus on behavior analysis in education: Achievements, challenges, and opportunities* (pp. 267–282), Upper Saddle River, NJ: Merrill/Prentice Hall.

Herr, S. S., O'Sullivan, J. L., & Dinerstein, R. D. (1999). Consent to extraordinary interventions. In R. D. Dinerstein, S. S. Herr, & J. L. O'Sullivan (Eds.), *A guide to consent* (pp. 111–122). Washington, DC: American Association on Intellectual and Developmental Disabilities.

Herrnstein, R. J. (1961). Relative and absolute strength of a response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior 4*, 267–272.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior 13*, 243–266.

Herrnstein, R. J., & Hineline, P. N. (1966). Negative reinforcement as shock-frequency reduction. *Journal of the Experimental Analysis of Behavior*, *9*, 421–430.

Herscovitch, B., Roscoe, E. M., Libby, M. E., Bourret, J. C., & Ahearn, W. H. (2009). A procedure for identifying precursors to problem behavior. *Journal of Applied Behavior Analysis*, *42*, 697–702. doi:10.1901/jaba.2009.42-697

Hersen, M., & Barlow, D. H. (1976). *Single case experimental designs: Strategies for studying behavior change*. New York: Pergamon Press.

Heward, L. D., Park, J., & Alber-Morgan, S. R. (2008). *The effects of indiscriminable contingencies on the completion and accuracy of independent academic work by second graders*. Paper presented at the 34th Annual Convention of the Association for Behavior Analysis International, Chicago, IL.

Heward, W. L. (1978, May). *The delayed multiple baseline design*. Paper presented at the Fourth Annual Convention of the Association for Behavior Analysis, Chicago.

Heward, W. L. (1978). Operant conditioning of a .300 hitter?: The effects of reinforcement on the offensive efficiency of a barnstorming baseball team. *Behavior Modification*, *2*, 25–40.

Heward, W. L. (1980). A formula for individualizing initial criteria for reinforcement. *Exceptional Teacher*, *1*(9), 7, 9.

Heward, W. L. (2003). Ten faulty notions about teaching and learning that hinder the effectiveness of special education. *The Journal of Special Education*, *36*(4), 186–205.

Heward, W. L. (2005). Reasons applied behavior analysis is good for education and why those reasons have been insufficient. In W. L. Heward, T. E. Heron, N. A. Neef, S. M. Peterson, D. M. Sainato, G. Cartledge, R. Gardner III, L. D. Peterson, S. B. Hersh, & J. C. Dardig (Eds.), *Focus on behavior analysis in education: Achievements, challenges, and opportunities* (pp. 316–348). Upper Saddle River, NJ: Merrill/Prentice Hall.

Heward, W. L. (2017). Why be a behavior analyst? In R. D. Holdsambeck & H. S. Pennypacker (Eds.), *Behavioral science: Tales of inspiration, discovery, and service. Omnibus edition* (pp. 359–383). Littleton, MA: The Cambridge Center for Behavioral Studies.

Heward, W. L., & Chance, P. (Guest Eds.). (2010). Special section: The human response to climate change: Ideas from behavior analysis. *The Behavior Analyst*, *33*, 145–206.

Heward, W. L., & Cooper, J. O. (1992). Radical behaviorism: A productive and needed philosophy for education. *Journal of Behavioral Education*, *2*, 345–365.

Heward, W. L., & Critchfield, T. S. (2019). *ABA from A-to-Z*. Manuscript in progress.

Heward, W. L., & Eachus, H. T. (1979). Acquisition of adjectives and adverbs in sentences written by hearing impaired and aphasic children. *Journal of Applied Behavior Analysis*, *12*, 391–400.

Heward, W. L., & Kimball, J. W. (2013, Spring/Summer). Sustaining sustainability with clueless contingencies. *Sustain Magazine*, *28*, 4–15.

Heward, W. L., & Silvestri, S. M. (2005). Antecedent. In G. Sugai & R. Horner (Eds.), *Encyclopedia of behavior modification and cognitive behavior therapy, Vol. 3: Educational applications* (pp. 1135–1137). Thousand Oaks, CA: Sage.

Heward, W. L., & Wood, C. L. (2003). Thursday afternoons with Don: Selections from three teleconference seminars on applied behavior analysis. In K. S. Budd & T. Stokes (Eds.), *A small matter of proof: The legacy of Donald M. Baer* (pp. 293–310). Reno, NV: Context Press.

Heward, W. L., with Alber-Morgan, S. R., & Konrad, M. (2017). *Exceptional children: An introduction to special education* (11th ed.). Upper Saddle River, NJ: Pearson.

Heward, W. L., Cooper, J. O., Heron, T. E., Gardner, R., III, & Sainato, D. M. (1995). Training leadership personnel for special education: The Ohio State University doctoral program in applied behavior analysis. *Teacher Education and Special Education*, *18*, 192–204.

Heward, W. L., Dardig, J. C., & Rossett, A. (1979). *Working with parents of handicapped children*. Upper Saddle River, NJ: Merrill/Prentice Hall.

Heward, W. L., Heron, T. E., Gardner, R., III, & Prayzer, R. (1991). Two strategies for improving students' writing skills. In G. Stoner, M. R. Shinn, & H. M. Walker (Eds.), *A school psychologist's interventions for regular education* (pp. 379–398). Washington, DC: National Association of School Psychologists.

Hewett, F. M. (1968). *The emotionally disturbed child in the classroom*. New York: McGraw-Hill.

Higbee, T. S., Carr, J. E., & Harrison, C. D. (1999). The effects of pictorial versus tangible stimuli in stimulus preference assessments. *Research in Developmental Disabilities*, *20*, 63–72.

Higbee, T. S., Carr, J. E., & Harrison, C. D. (2000). Further evaluation of the multiple-stimulus preference assessment. *Research in Developmental Disabilities*, *21*, 61–73.

Higgins, J. W., Williams, R., & McLaughlin, T. F. (2001). The effects of a token economy employing instructional consequences for a third-grade student with learning disabilities: A data-based case study. *Education and Treatment of Children*, *24*(1), 99–106.

Higgins, W. J., Luczynski, K. C., Carroll, R. A., Fisher, W. W., & Mudford, O. C. (2017). Evaluation of a Telehealth training package to remotely train staff to conduct a preference assessment. *Journal of Applied Behavior Analysis*, *50*, 238–251.

Himle, M. B., Miltenberger, R. G., Gatheridge, B., & Flessner, C., (2004). An evaluation of two procedures for training skills to prevent gun play in children. *Pediatrics*, *113*, 70–77.

Hine, J. F., Ardoin, S. P., & Foster, T. E. (2015). Decreasing transition times in elementary school classrooms: Using computer-assisted instruction to automate intervention components. *Journal of Applied Behavior Analysis*, *48*, 495–510.

Hineline, P. N. (2005). The aesthetics of behavioral arrangements. *The Behavior Analyst*, *28*, 15-28.

Hineline, P. N. (2018). Narrative: How it works and why it's important. *Perspectives on Behavior Science*. Advance online publication.

Hineline, P. N., & Rosales-Ruiz, J. (2013). Behavior in relation to aversive events: Punishment and negative reinforcement. In G. J. Madden (Ed.), *APA handbook of behavior analysis, Vol. 1: Methods and principles* (pp. 483–512). Washington, DC: American Psychological Association.

Hines, T. M. (2001). The Doman-Delacato patterning treatment for brain damage. *The Scientific Review of Alternative Medicine*, *5*(2), 80–89.

Hirst, E. S. J., Dozier, C. L., & Payne, S. W. (2016). Efficacy of and preference for reinforcement and response cost in token economies. *Journal of Applied Behavior Analysis*, *49*, 329–345.

Hitchcock, J. H., Kratochwill, T. R., & Chezan, L. C. (2015). What Works Clearinghouse standards and generalization of single-case design evidence. *Journal of Behavioral Education*, *24*, 459–469. doi 10.1007/s10864-015-9224-1

Hixson, M. D. (2004). Behavioral cusps, basic behavioral repertoires, and cumulative-hierarchical learning. *The Psychological Record*, *54*, 387–403.

Hixson, M. D., Reynolds, J. L., Bradley-Johnson, S., & Johnson, C. M. (2011). Cumulative hierarchical learning and behavioral cusps. In E. A. Mayhill & J. A. Mulick (Eds.), *Behavioral foundations of effective autism treatment* (pp. 137–154). Cambridge, MA: Sloan Publishing.

Hobbs, T. R., & Holt, M. M. (1976). The effects of token reinforcement on the behavior of delinquents in cottage settings. *Journal of Applied Behavior Analysis*, *9*, 189–198.

Hoch, H., McComas, J. J., Johnson, L, Faranda, N., & Guenther, S. L. (2002). The effects of magnitude and quality of reinforcement on choice responding during play activities. *Journal of Applied Behavior Analysis*, *35*, 171–181.

Hoch, H., McComas, J. J., Thompson, A. L., & Paone, D. (2002). Concurrent reinforcement schedules: Behavior change and maintenance without extinction. *Journal of Applied Behavior Analysis*, *35*, 155–169.

Hockett, C. F. (1960). The origin of speech. *Scientific American*, *203*, 88–96.

Hodos, W. (1961) Progressive ratio as a measure of reward strength. *Science*, *6*, 387–394.

Hoffman, K., & Falcomata, T. S. (2014). An evaluation of resurgence of appropriate communication in individuals with autism who exhibit severe problem behavior. *Journal of Applied Behavior Analysis*, *47*, 651–656. doi:10.1002/jaba.144

Hogan, D. E., & Zentall, T. R. (1977). Backward associations in the pigeon. *American Journal of Psychology*, *90*, 3–15.

Holcombe, A., Wolery, M., Werts, M. G., & Hrenkevich, P. (1993). Effects of instructive feedback on future learning. *Journal of Behavioral Education*, *3*, 259–285.

Holifield, C., Goodman, J., Hazelkorn, M., & Heflin, L. J. (2010). Using self-monitoring to increase attending to task and academic accuracy in children with autism. *Focus on Autism and Other Developmental Disabilities*, *25*(4), 230–238. doi:10.1002/jaba.377

Holland, J. G. (1978). Behaviorism: Part of the problem or part of the solution? *Journal of Applied Behavior Analysis*, *11*, 163–174.

Holland, J. G., & Skinner, B. F. (1961). *The analysis of behavior: A program for self-instruction*. New York: McGraw-Hill.

Holman, J., Goetz, E. M., & Baer, D. M. (1977). The training of creativity as an operant and an examination of its generalization characteristics. In B. C. Etzel, J. M. LeBlanc, & D. M. Baer (Eds.), *New developments in behavioral research: Theory, method, and practice* (pp. 441–471). Hillsdale, NJ: Erlbaum.

Holmes, G., Cautela, J., Simpson, M., Motes, P., & Gold, J. (1998). Factor structure of the School Reinforcement Survey Schedule: School is more than grades. *Journal of Behavioral Education*, *8*, 131–140.

Holth, P. (2003). Generalized imitation and generalized matching to sample. *The Behavior Analyst*, *26*, 155–158.

Holth, P. (2012). Variability as an operant? *The Behavior Analyst*, *35*, 243–248. doi:10.1007/s40614-017-0083-z

Holth, P. (2017). Multiple-exemplar training: Some strengths and limitations. *The Behavior Analyst*, *40*, 225–241. doi:10.1007/s40614-017-0083-z

Holz, W. C., & Azrin, N. H. (1961). Discriminative properties of punishment. *Journal of the Experimental Analysis of Behavior*, *4*, 225–232. doi:10.1901/jeab.1961.4-225

Holz, W. C., & Azrin, N. H. (1962). Recovery during punishment by intense noise. *Journal of the Experimental Analysis of Behavior*, *6*, 407–412.

Holz, W. C., Azrin, H. N., & Ayllon, T. (1963). Elimination of behavior of mental patients by response-produced extinction. *Journal of the Experimental Analysis of Behavior*, *6*, 407–412.

Homme, L., Csanyi, A. P., Gonzales, M. A., & Rechs, J. R. (1970). *How to use contingency contracting in the classroom*. Champaign, IL: Research Press.

Honig, W. K. (Ed.). (1966). *Operant behavior: Areas of research and application*. New York: Appleton-Century-Crofts.

Honnen, T. J., & Kleinke, C. L. (1990). Prompting bar patrons with signs to take free condoms. *Journal of Applied Behavior Analysis*, *23*, 215–217.

Hoon, A., Dymond, S., Jackson, J. W., & Dixon, M. R. (2007). Manipulating contextual control over simulated slot machine gambling. *Analysis of Gambling Behavior*, *1*, 109–122.

Hoon, A., Dymond, S., Jackson, J. W., & Dixon, M. R. (2008). Contextual control of slot-machine gambling: Replication and extension. *Journal of Applied Behavior Analysis*, *41*, 467–470.

Hoover, H. D., Dunbar, S. B., & Frisbie, D. A. (2007). *Iowa Tests of Basic Skills*. Rolling Meadows, IL: Riverside.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, *65*, 185–241.

Horne, P. J., & Lowe, C. F. (2000). Putting the naming account to the test: Preview of an experimental program. In J. C. Leslie & D. Blackman (Eds.), *Experimental and applied analysis of human behavior*. Reno, NV: Context Press.

Horne, P. J., Hardman, C. A., Lowe, C. F., Tapper, K., Le Noury, J., Madden, P., & Doody, M. (2009). Increasing parental provision and children's consumption of lunchbox fruit and vegetables in Ireland: The Food Dudes intervention. *European Journal of Clinical Nutrition*, *63*, 613–618.

Horne, P. J., Hughes, J. C., & Lowe, C. F. (2006). Naming and categorization in young children: IV: Listener behavior training and transfer for function. *Journal of the Experimental Analysis of Behavior*, *85*, 247–273.

Horne, P. J., Lowe, C. F., & Randle, V. R. L. (2004). Naming and categorization in young children: II. Listener behavior training. *Journal of the Experimental Analysis of Behavior*, *81*, 267–288.

Horne, P. J., Tapper, K., Lowe, C. F., Hardman, C. A., Jackson, M. C., & Woolner, J. (2004). Increasing children's fruit and vegetable consumption: A peer-modeling and rewards-based intervention. *European Journal of Clinical Nutrition*, *58*, 1649–1660.

Horner, R., Carr, E., Strain, P., Todd, A., & Reed, H. (2002). Problem behavior interventions for young children with autism: A research synthesis. *Journal of Autism and Developmental Disorders*, *32*, 423–446.

Horner, R., Sugai, G., Smolkowski, K., Todd, A., Nakasato, J., & Esperanza, J. (2009). A randomized control trial of school-wide positive behavior support in elementary schools. *Journal of Positive Behavior Interventions*, *11*(3), 113–144.

Horner, R. D. (1980). The effects of an environmental "enrichment" program on the behavior of institutionalized profoundly retarded children. *Journal of Applied Behavior Analysis*, *13*, 473–491.

Horner, R. D., & Baer, D. M. (1978). Multiple-probe technique: A variation on the multiple baseline design. *Journal of Applied Behavior Analysis*, *11*, 189–196.

Horner, R. H. (1990). Ideology, technology and typical community settings: The use of severe aversive stimuli. *American Journal of Mental Retardation*, *28*(4), 166–168.

Horner, R. H. (2002). On the status of knowledge for using punishment: A commentary. *Journal of Applied Behavior Analysis, 35*, 465–467.

Horner, R. H., & Kratochwill, T. R. (2012). Synthesizing single-case research to identify evidence-based practices: Some brief reflections. *Journal of Behavioral Education, 21*, 266–272. doi:10.1007/s10864-012-9152-2

Horner, R. H., & McDonald, R. S. (1982). Comparison of single instance and general case instruction in teaching of a generalized vocational skill. *Journal of the Association for Persons with Severe Handicaps, 7*(3), 7–20.

Horner, R. H., Day, H. M., & Day, J. R. (1997). Using neutralizing routines to reduce problem behaviors. *Journal of Applied Behavior Analysis, 30*, 601–614.

Horner, R. H., Day, M., Sprague, J., O'Brien, M., & Heathfield, L. (1991). Interspersed requests: A nonaversive procedure for reducing aggression and self-injury during instruction. *Journal of Applied Behavior Analysis, 24*, 265–278.

Horner, R. H., Dunlap, G., & Koegel, R. L. (1988). *Generalization and maintenance: Life-style changes in applied settings*. Baltimore: Paul H. Brookes.

Horner, R. H., Eberhard, J. M., & Sheehan, M. R. (1986). Teaching generalized table bussing: The importance of negative teaching examples. *Behavior Modification, 10*, 457–471.

Horner, R. H., Sprague, J., & Wilcox, B. (1982). Constructing general case programs for community activities. In B. Wilcox & G. T. Bellamy (Eds.), *Design of high school programs for severely handicapped students* (pp. 61–98). Baltimore: Paul H. Brookes.

Horner, R. H., Williams, J. A., & Steveley, J. D. (1987). Acquisition of generalized telephone use by students with moderate and severe mental retardation. *Research in Developmental Disabilities, 8*, 229–247.

Hosp, M. K., & Hosp, J. L. (2003). Curriculum-based measurement for reading, spelling, and math: How to do it and why. *Preventing School Failure, 48*(1), 10–17.

Hott, B. L., & Brigham, F. (2018). Effects of response options on the mathematics performance of secondary students with emotional or behavioral disorders. *Exceptionality*, published online 20 June 2018. https://doi.org/10.1080/09362835.2018.1480950

House Rich, S. E., & Duhon, G. J. (2014). Using brief academic assessments to determine generalization strategies. *Journal of Behavioral Education, 23*, 401–420. doi:10.1007/s10864-014-9212-x

Howard, J. S., & Rice, D. S. (1988). Establishing a generalized autoclitic repertoire in preschool children. *The Analysis of Verbal Behavior, 6*, 45–59.

Howard, J. S., Sparkman, C. R., Cohen, H. G., Green, G., & Stanislaw, H. (2005). A comparison of intensive behavior analytic and eclectic treatments for young children with autism. *Research in Developmental Disabilities, 26*, 359–383.

Howard, V. J., & DiGennaro Reed, F. D. (2014). Training shelter volunteers to teach dog compliance. *Journal of Applied Behavior Analysis, 47*, 344–359. doi:10.1002/jaba.120

Howard, V. J., & Digennaro Reed, F. D. (2015). An evaluation of training procedures for animal shelter volunteers. *Journal of Organizational Behavior Management, 35*, 296–320.

Hübner, M. M. C., Austin, J., & Miguel, C. F. (2008). The effects of praising qualifying autoclitics on the frequency of reading. *The Analysis of Verbal Behavior, 24*, 55–62.

Hübner, M. M. C., Gomes, R. C., & McIlvane, W. J. (2009). Recombinative generalization in minimal verbal unit-based reading instruction for pre-reading children. *Experimental Analysis of Human Behavior Bulletin, 27*, 11–17.

Hughes, C. (1992). Teaching self-instruction using multiple exemplars to produce generalized problem-solving by individuals with severe mental retardation. *Journal on Mental Retardation, 97*, 302–314.

Hughes, C. (1997). Self-instruction. In M. Agran (Ed.), *Self-directed learning: Teaching self-determination skills* (pp. 144–170). Pacific Grove, CA: Brooks/Cole.

Hughes, C., & Lloyd, J. W. (1993). An analysis of self-management. *Journal of Behavioral Education, 3*, 405–424.

Hughes, C., & Rusch, F. R. (1989). Teaching supported employees with severe mental retardation to solve problems. *Journal of Applied Behavior Analysis, 22*, 365–372.

Hughes, C., Harmer, M. L., Killian, D. J., & Niarhos, F. (1995). The effects of multiple-exemplar self-instructional training on high school students' generalized conversational interactions. *Journal of Applied Behavior Analysis, 28*, 201–218.

Hughes, J. (2006). A pilot study of naturally occurring high-probability request sequences in hostage negotiations. *Journal of Applied Behavior Analysis, 42*, 491–496.

Hughes, S. J., & Barnes-Holmes, D. (2013). A functional approach to the study of implicit cognition: The IRAP and the REC model. In S. Dymond & B. Roche (Eds.), *Advances in relational frame theory and contextual behavioural science: Research & applications* (pp. 97–126). Oakland, CA: New Harbinger.

Hughes, T. A., & Fredrick, L. D. (2006). Teaching vocabulary to students with learning disabilities using classwide peer tutoring and constant time delay. *Journal of Behavioral Education, 15*, 1–23.

Huitema, B., Van Houten, R., & Manal, H. (2014). Time-series intervention analysis of pedestrian countdown timer effects. *Accident Analysis and Prevention, 72*, 23–31.

Hull, C. L. (1943). *Principles of behavior*. New York: Appleton-Century-Crofts.

Hume, K. M., & Crossman, J. (1992). Musical reinforcement of practice behaviors among competitive swimmers. *Journal of Applied Behavior Analysis, 25*, 665–670.

Hume, K. M., Martin, G. L., Gonzalez, P., Cracklen, C., & Genthon, S. (1985). A self-monitoring feedback package for improving freestyle figure skating practice. *Journal of Sport Psychology, 7*, 333–345.

Humphrey, L. L., Karoly, P., & Kirschenbaum, D. S. (1978). Self-management in the classroom: Self-imposed response cost versus self-reward. *Behavior Therapy, 9*, 592–601.

Hundert, J., & Batstone, D. (1978). A practical procedure to maintain pupils' accurate self-rating in a classroom token program. *Behavior Modification, 2*, 93–112. doi.org/10.1177/014544557821006

Hundert, J., & Bucher, B. (1978). Pupils' self-scored arithmetic performance: A practical procedure for maintaining accuracy. *Journal of Applied Behavior Analysis, 11*, 304.

Hunt, P., & Goetz, L. (1988). Teaching spontaneous communication in natural settings using interrupted behavior chains. *Topics in Language Disorders, 9*, 58–71.

Hunt, P., Soto, G., Maier, J., & Doering, K. (2003). Collaborative learning to support students at risk and students with severe disabilities in general education classrooms. *Exceptional Children, 69*, 315–332. doi:10.1177/001440290306900304

Hunter, M., & Rosales-Ruiz, J. (2018). *An introduction to PORTL: The Portable Operant Research and Teaching Lab*. Retrieved January 2, 2018, from https://www.artandscienceofanimaltraining.org/tools/portl-shaping-game/

Hunter, S. H., Beverly, M., Parkinson, J., & Hughes, J. C. (2016). Increasing high school students' maths skills with the use of SAFMEDS class-wide. *European Journal of Behavior Analysis, 18*, 192–199.

Hunter, W. C., Maheady, L., Jasper, A. D., Williamson, R. L., Murley, R. C., & Stratton, E. (2015). Numbered Heads Together as a Tier I instructional strategy in multitiered systems of support. *Education and Treatment of Children, 38*(3), 345–362.

Hurley, A. D. N., & O'Sullivan, J. L. (1999). Informed consent for health care. In R. D. Dinerstein, S. S. Herr, & J. L. O'Sullivan (Eds.), *A guide to consent* (pp. 39–55). Washington, DC: American Association on Intellectual and Developmental Disabilities.

Hursh, S. R., Madden, G. J., Spiga, R., DeLeon, I. G., & Francisco, M. T. (2013). The translational utility of behavioral economics: The experimental analysis of consumption and choice. In G. J. Madden (Ed.), *APA handbook of behavior analysis, Vol. 2: Translating principles into practice* (pp. 3191–3224). Washington, DC: American Psychological Association.

Hutchinson, R. R. (1977). By-products of aversive control. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 415–431). Upper Saddle River, NJ: Prentice Hall.

Imam, A. A. (2006). Experimental control of nodality via equal presentations of control discriminations in different equivalence protocols under speed and no-speed conditions. *Journal of the Experimental Analysis of Behavior, 85*, 107–124.

Ingram, R. E. (1990). Self-focused attention in clinical disorders: Review and a conceptual model. *Psychological Bulletin, 107*, 156–176.

Ingvarsson, E. T., Cammilleri, A. P., & Macias, H. (2012). Emergent listener responses following intraverbal training in children with autism. *Research in Autism Spectrum Disorders*, *6*, 654–664. doi.org/10.1016/j.rasd.2011.09.009

Ingvarsson, E. T., Kramer, R. L., Carp, C. L., Pétursdóttir, A. I., & Macias, H. (2016). Evaluation of a blocked-trials procedure to establish complex stimulus control over intraverbal responses in children with autism. *The Analysis of Verbal Behavior*, doi 10.1007/s40616-016-0071-5

Ingvarsson, E. T., Tiger, J. H., Hanley, G. P., & Stephenson, K. M. (2007). An evaluation of intraverbal training to generate socially appropriate responses to novel questions. *Journal of Applied Behavior Analysis*, *40*, 411–429.

Interactive Autism Network. (2019) *Gold standard of evidence: The randomized controlled trial (RCT)*. Retrieved January 8, 2019, at https://iancommunity.org/cs/understanding_research/randomized_controlled_trials

Intergovernmental Panel on Climate Change (IPCC). (2014). *Climate change 2014: Synthesis report*. Geneva, Switzerland: Intergovernmental Panel on Climate Change. Available at https://ipcc.ch/report/ar5/

Intergovernmental Panel on Climate Change (IPCC). (2018). *Global warming of 1.5 °C*. Incheon, Korea: Intergovernmental Panel on Climate Change. Available at https://www.ipcc.ch/sr15/

Irvin, D. S., Thompson, T. J., Turner, W. D., & Williams, D. E. (1998). Utilizing response effort to reduce chronic hand mouthing. *Journal of Applied Behavior Analysis*, *31*, 375–385.

Isaacs, W., Thomas, I., & Goldiamond, I. (1960). Application of operant conditioning to reinstate verbal behavior in psychotics. *Journal of Speech and Hearing Disorders*, *25*, 8–12.

Isaksen, J., and Holth, P. (2009). An operant approach to teaching joint attention skills to children with autism. *Behavioral Interventions*, *24*, 215–236.

Iverson, I. H. (2013). Single-case research methods: An overview. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. 3–32). Washington, DC: American Psychological Association.

Ivy, J. W., Meindl, J. N., Overly, E., & Robson, K. M. (2017). Token economy: A systematic review of procedural descriptions. *Behavior Modification*, *41*(5), 708–737. doi:10.1177/0145445517699559

Ivy, J. W., Neef, N. A., Meindl, J. N., & Miller, N. (2016). A preliminary examination of motivating operations and reinforcer class interaction. *Behavior Interventions*, *31*, 180–194.

Iwata, B. A. (1987). Negative reinforcement in applied behavior analysis: An emerging technology. *Journal of Applied Behavior Analysis*, *20*, 361–378.

Iwata, B. A. (1988). The development and adoption of controversial default technologies. *Behavior Analyst*, *11*(2), 149–157.

Iwata, B. A. (1991). Applied behavior analysis as technological science. *Journal of Applied Behavior Analysis*, *24*, 421–424.

Iwata, B. A. (2006). On the distinction between positive and negative reinforcement. *The Behavior Analyst*, *29*, 121–123.

Iwata, B. A., & DeLeon, I. (1996). *The functional analysis screening tool*. Gainesville, FL: The Florida Center on Self-Injury, The University of Florida.

Iwata, B. A., & Michael, J. L. (1994). Applied implications of theory and research on the nature of reinforcement. *Journal of Applied Behavior Analysis*, *27*, 183–193.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994a). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis*, *27*, 197–209. doi:10.1901/jaba.1994.27-197

Iwata, B. A., Pace, G. M., Cowdery, G. E., & Miltenberger, R. G. (1994). What makes extinction work: An analysis of procedural form and function. *Journal of Applied Behavior Analysis*, *27*, 131–144.

Iwata, B. A., Pace, G. M., Dorsey, M. F., Zarcone, J. R., Vollmer, T. R., Smith, R. G., & Willis, K. D. (1994b). The functions of self-injurious behavior: An experimental-epidemiological analysis. *Journal of Applied Behavior Analysis*, *27*, 215–240.

Iwata, B. A., Pace, G. M., Kissel, R. C., Nau, P. A., & Farber, J. M. (1990). The Self-Injury Trauma (SIT) Scale: A method for quantifying surface tissue damage caused by self-injurious behavior. *Journal of Applied Behavior Analysis*, *23*, 99–110.

Iwata, B. A., Rolider, N. U., & Dozier, C. L. (2009). Evaluation of timeout programs through phased withdrawal. *Journal of Applied Research in Intellectual Disabilities*, *22*, 203–209.

Iwata, B. A., Smith, R. G., & Michael, J. (2000). Current research on the influence of establishing operations on behavior in applied settings. *Journal of Applied Behavior Analysis*, *33*(4), 411–418.

Jackson, M. L., Mendoza, D. R., & Adams, A. N. (2014). Teaching a deictic relational repertoire to children with autism. *The Psychological Record*, *64*, 791–802.

Jacobs, E. A., Borrero, J. C., & Vollmer, T. R. (2013). Translational applications of quantitative choice modals. In G. J. Madden (Ed.), *APA handbook of behavior analysis, Vol. 2: Translating principles into practice* (pp. 387–404). Washington, DC: American Psychological Association.

Jacobs, H. E., Fairbanks, D., Poche, C. E., & Bailey, J. S. (1982). Multiple incentives in encouraging car pool formation on a university campus. *Journal of Applied Behavior Analysis*, *15*, 141–149.

Jacobson, J. W., Foxx, R. M., & Mulick, J. A. (Eds.). (2005). *Controversial therapies for developmental disabilities: Fads, fashion, and science in professional practice*. Hillsdale, NJ: Erlbaum.

Jacobsen, J. W., Foxx, R. M., & Mulick, J. A. (2016). Facilitated communication: The ultimate fad treatment. In R. M. Foxx & J. A. Mulick (Eds.), *Controversial therapies for autism and intellectual disabilities: Fads, fashion, and science in professional practice* (2nd ed.) (pp. 136–153). New York: Routledge.

Jacobson, J. M., Bushell, D., & Risley, T. (1969). Switching requirements in a Head Start classroom. *Journal of Applied Behavior Analysis*, *2*, 43–47.

Jahr, E., & Eldevik, S. (2002). Teaching cooperative play to typical children utilizing a behavior modeling approach: A systematic replication. *Behavioral Interventions*, *17*, 145–157.

James, J. E. (1981). Behavioral self-control of stuttering using time-out from speaking. *Journal of Applied Behavior Analysis*, *14*, 25–37.

Jarmolowizc, D. P., & Lattal, K. A. (2010). On distinguishing progressively increasing response requirements for reinforcement. *The Behavior Analyst*, *33*, 119–125.

Jarvis, B. P., & Dallery, J. (2017). Internet-based self-tailored deposit contracts to promote smoking reduction and abstinence. *Journal of Applied Behavior Analysis*, *50*, 189–205.

Jason, L. A., & Liotta, R. F. (1982). Reduction of cigarette smoking in a university cafeteria. *Journal of Applied Behavior Analysis*, *15*, 573–577.

Jeffries, T., Crosland, K., & Miltenberger, R. (2016). Evaluating a tablet application and differential reinforcement to increase eye contact in children with autism. *Journal of Applied Behavior Analysis*, *49*, 182–187.

Jennings, A. M., & Miguel, C. F. (2017). Training intraverbal bidirectional naming to establish generalized equivalence class performances. *Journal of the Experimental Analysis of Behavior*, *108*, 269–289. doi:10.1002/jeab.277

Jerome, J., & Sturmey, P. (2008). Reinforcing efficacy of interactions with preferred and nonpreferred staff under progressive-ratio schedules. *Journal of Applied Behavior Analysis*, *41*, 221–225.

Jessel, J., & Borrero, J. C. (2014). A laboratory comparison of two variations of differential-reinforcement-of-low-rate procedures. *Journal of Applied Behavior Analysis*, *47*, 314–324. doi:10.1002/jaba.114

Jessel, J., & Ingvarsson, E. T. (2016). Recent advances in applied research on DRO procedures. *Journal of Applied Behavior Analysis*, *49*, 991–995. doi:10.1002/jaba.323

Jessel, J., Borrero, J. C., & Becraft, J. L. (2015). Differential reinforcement of other behavior increases untargeted behavior. *Journal of Applied Behavior Analysis*, *48*, 402–416. doi:10.1002/jaba.204

Jessel, J., Hanley, G. P., & Ghaemmaghami, M. (2016). Interview-informed synthesized contingency analyses: Thirty replications and reanalysis. *Journal of Applied Behavior Analysis*, *49*, 576–595. doi:10.1002/jaba.316

Jezierski, T., Adamkiewicz, E., Walczak, M., Sobczynska, M., Gorecka-Bruzda-A., Ensminger, J., & Papet, E. (2014). Efficacy of drug detection by fully-trained police dogs varies by breed, training level, type of drug, and search environment. *Forensic Science International*, *237*, 112–118.

Jimenez, B. A., Browder, D. M., Spooner, F., & Dibiase, W. (2012). Inclusive inquiry science using peer-mediated embedded instruction for students with moderate intellectual disabilities. *Exceptional Children*, *78*(3), 301–317.

Johnson, B. M., Miltenberger, R. G., Knudson, P., Egemo-Helm, K., Kelso, P., Jostad, C., & Langley, L. (2006). A preliminary evaluation of two behavioral skills training procedures for teaching abduction-prevention skills to school children. *Journal of Applied Behavior Analysis*, *39*, 25–34.

Johnson, C., Meleshkevich, O., & Dube, W. V. (2014). Merging separately established stimulus classes with outcome-specific reinforcement. *Journal of the Experimental Analysis of Behavior*, *101*, 38–50.

Johnson, K., & Street, E. M. (2014). Precision teaching: The legacy of Ogden Lindsley. In F. K. McSweeney & E. S. Murphy (Eds.) (2014). *The Wiley-Blackwell handbook of operant and classical conditioning* (pp. 581–609). Oxford, UK: John Wiley & Sons.

Johnson, K. R., & Layng, T. V. J. (1992). Breaking the structuralist barrier: Literacy and numeracy with fluency. *American Psychologist*, *47*, 1475–1490.

Johnson, K. R., & Layng, T. V. J. (1994). The Morningside model of generative instruction. In R. Gardner III, D. M. Sainato, J. O. Cooper, T. E. Heron, W. L. Heward, J. Eshleman, & T. A. Grossi (Eds.), *Behavior analysis in education: Focus on measurably superior instruction* (pp. 173–197). Pacific Grove, CA: Brooks/Cole.

Johnson, M. W., & Bruner, N. R. (2012). The sexual discounting task: HIV risk behavior and the discounting of delayed sexual rewards in cocaine dependence. *Drug and Alcohol Dependence*, *123*, 15–21.

Johnson, P. E., Jr., Perrin, C. J., Salo, A., Deschaine, E., & Johnson, B. (2016). Use of an explicit rule decreases procrastination in university students. *Journal of Applied Behavior Analysis*, *49*, 346–358.

Johnston, J. M. (1979). On the relation between generalization and generality. *The Behavior Analyst*, *2*, 1–6.

Johnston, J. M. (1991). We need a new model of technology. *Journal of Applied Behavior Analysis*, *24*, 425–427.

Johnston, J. M. (2013). *Radical behaviorism for ABA practitioners*. Cornwall-on-Hudson, NY: Sloan.

Johnston, J. M., & Pennypacker, H. S. (1980). *Strategies and tactics of human behavioral research*. Hillsdale, NJ: Erlbaum.

Johnston, J. M., & Pennypacker, H. S. (1993a). *Strategies and tactics of human behavioral research* (2nd ed.). Hillsdale, NJ: Erlbaum.

Johnston, J. M., & Pennypacker, H. S. (1993b). *Readings for strategies and tactics of behavioral research* (2nd ed.). Hillsdale, NJ: Erlbaum.

Johnston, J. M., & Pennypacker, H. S. (2009). *Strategies and tactics of behavioral research* (3rd ed.). New York: Routledge.

Johnston, J. M., Mellichamp, F. H., Shook, G. L., & Carr, J. E. (2014). Determining BACB examination content and standards. *Behavior Analysis in Practice*, *7*, 3–9. doi 10.1007/s40617-014-0003-6

Johnston, M. K., Kelly, C. S., Harris, F. R., & Wolf, M. M. (1966). An application of reinforcement principles to the development of motor skills of a young child. *Child Development*, *37*, 370–387.

Johnson, T. (1973). *Addition and subtraction math program with stimulus shaping and stimulus fading*. Produced pursuant to a grant from the Ohio Department of Education, BEH Act, P.L. 91-203, Title VI-G; OE G-0-714438(604). J. E. Fisher & J. O. Cooper, project codirectors.

Jones, E. A., Carr, E. G., & Feeley, K. M. (2006). Multiple effects of joint attention intervention for children with autism. *Behavior Modification*, *30*, 782–834.

Jones, F. H., & Miller, W. H. (1974). The effective use of negative attention for reducing group disruption in special elementary school classrooms. *Psychological Record*, *24*, 435–448.

Jones, F. H., Fremouw, W., & Carples, S. (1977). Pyramid training of elementary school teachers to use a classroom management "skill package." *Journal of Applied Behavior Analysis*, *10*, 239–254.

Jones, R. R., Vaught, R. S., & Weinrott, M. (1977). Time-series analysis in operant research. *Journal of Applied Behavior Analysis*, *10*, 151–166.

Jones, R. T., Nelson, R. E., & Kazdin, A. E. (1977). The role of external variables in self-reinforcement. *Behavior Modification*, *1*, 147–178.

Jones, V. F., & Jones, L. S. (1981). *Responsible classroom discipline: Creating positive learning environments and solving problems*. Boston: Allyn & Bacon.

Joseph, L. M., & Konrad, M. (2009). Twenty ways to help students self-manage their academic performance. *Intervention in School and Clinic*, *44*, 246–249.

Jowett Hirst, E. S., Dozier, C. L., & Payne, S. W. (2016). Efficacy of and preference for reinforcement and response cost in token economies. *Journal of Applied Behavior Analysis*, *49*, 329–345. doi:10.1002/jaba.294

Joyce, B. G., Joyce, J. H., & Wellington, B. (1993). Using stimulus equivalence procedures to teach relationships between English and Spanish words. *Education and Treatment of Children*, *16*, 48–65.

Jung, S., Sainato, D. M., & Davis, C. A. (2008). Using high-probability request sequences to increase social interactions in young children with autism. *Journal of Early Intervention*, *30*, 163–187. doi:10.1177/1053815108317970

Kabler, M. L. (1976). *Teaching fourth grade children to use self-contracting as a form of self control* (Doctoral dissertation). Retrieved from The Ohio State University, ProQuest Dissertations Publishing, 7702427.

Kachanoff, R., Leveille, R., McLelland, H., & Wayner, M. J. (1973). Schedule induced behavior in humans. *Physiology and Behavior*, *11*, 395–398.

Kagan, S., & Kagan, M. (2009). *Kagan cooperative learning*. San Clemente, CA: Kagan Publishing.

Kahng, S., & Iwata, B. A. (1999). Correspondence between outcomes of brief and extended functional analyses. *Journal of Applied Behavior Analysis*, *32*, 149–160. doi:10.1901/jaba.1999.32-149

Kahng, S., Chung, K.-M., Gutshall, K., Pitts, S. C., Kao, J., & Girolami, K. (2010). Consistent visual analyses of intrasubject data. *Journal of Applied Behavior Analysis*, *43*, 35–45. doi:10.1901/jaba.2010.43-35

Kahng, S., Hausman, N. L., Fisher, A. B., Donaldson, J. M., Cox, J. R., Lugo, M., & Wiskow, K. M. (2015). The safety of functional analyses of self-injurious behavior. *Journal of Applied Behavior Analysis*, *48*, 107–114.

Kahng, S., Iwata, B. A., Thompson, R. T., & Hanley, G. P. (2000). A method for identifying satiation versus extinction effects under noncontingent reinforcement schedules. *Journal of Applied Behavior Analysis*, *33*, 419–432.

Kahng, S. W., & Iwata, B. A. (1998). Computerized systems for collecting real-time observational data. *Journal of Applied Behavior Analysis*, *31*, 253–261.

Kahng, S. W., & Iwata, B. A. (1999). Correspondence between outcomes of brief and extended functional analyses. *Journal of Applied Behavior Analysis*, *32*, 149–159.

Kahng, S. W., & Iwata, B. A. (2000). Computer systems for collecting real-time observational data. In T. Thompson, D. Felce, & F. J. Symons (Eds.), *Behavioral observation: Technology and applications in developmental disabilities* (pp. 35–45). Baltimore: Paul H. Brookes.

Kahng, S. W., Ingvarsson, E. T., Quigg, A. M., Seckinger, K. E., & Teichman, H. M. (2011). Defining and measuring behavior. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of Applied Behavior Analysis* (pp.113–131). New York: Guilford Press.

Kahng, S. W., Iwata, B. A., DeLeon, I. G., & Wallace, M. D. (2000). A comparison of procedures for programming noncontingent reinforcement schedules. *Journal of Applied Behavior Analysis*, *33*, 223–231.

Kahng, S. W., Iwata, B. A., Fischer, S. M., Page, T. J., Treadwell, K. R. H., Williams, D. E., & Smith, R. G. (1998). Temporal distributions of problems behavior based on scatter plot analysis. *Journal of Applied Behavior Analysis*, *31*, 593–604.

Kahng, S. W., Iwata, B. A., Thompson, R. H., & Hanley, G. P. (2000). A method for identifying satiation versus extinction effects under noncontingent reinforcement schedules. *Journal of Applied Behavior Analysis*, *33*, 419–432.

Kahng, S. W., Tarbox, J., & Wilke, A. (2001). Use of multicomponent treatment for food refusal. *Journal of Applied Behavior Analysis*, *34*, 93–96.

Kaimowitz v. Department of Mental Health. Civil No. 73-19434-AW (Cir. Ct. of Wayne County, Mich. Filed July 10, 1973). In *Mental Disability Law Reporter*, *1*(2), September-October 1976 (pp. 147–154). American Bar Association.

Kame'enui, E. (2002, September). *Beginning reading failure and the quantification of risk: Behavior as the supreme index*. Presentation to the Focus on Behavior Analysis in Education Conference, Columbus, OH.

Kamin, L. J. (1957). The gradient of delay of secondary reward of avoidance learning tested on avoidance trials only. *Journal of Comparative and Physiological Psychology*, *50*, 450–456.

Kamps, D., Conklin, C., & Wills, H. (2015). Use of self-management with the CW-FIT group contingency program. *Education and Treatment of Children*, *38*, 1–32.

Kanfer, F. H. (1976). *The many faces of self-control, or behavior modification changes its focus*. Paper presented at the Fifth International Banff Conference, Banff, Alberta, Canada.

Kanfer, F. H., & Karoly, P. (1972). Self-control: A behavioristic excursion into the lion's den. *Behavior Therapy*, *3*, 398–416.

Kang, S., O'Reilly, M., Lancioni, G., Falcomata, T. S., Sigafoos, J., & Xu, Z. (2013). Comparison of the predictive validity and consistency among preference assessment procedures: A review of the literature. *Research in Developmental Disabilities*, *34*, 1125–1133. https://doi.org/10.1016/j.ridd.2012.12.021

Kangas, B. D., & Cassidy, R. N. (2010). Requiem for my lovely. *Journal of the Experimental Analysis of Behavior*, *95*, 269. doi:10.1901/jeab.2011.95-269

Kanter, J. W., Baruch, D. E., & Gaynor, S. T. (2006). Acceptance and commitment therapy and behavioral activation for the treatment of depression: Description and comparison. *The Behavior Analyst*, *29*, 161–185.

Kantor, J. R. (1958). Interbehavioral psychology. Fisher, R.A. (1925). *Statistical methods for research workers*. London, UK: Oliver & Boyd.

Kantor, J. R. (1959). *Interbehavioral psychology: A sample of scientific system construction* (Rev. ed.). Granville, OH: Principia Press.

Kaplan, B. A., Reed, D. D., & Jarmolowicz, D. P. (2016). Effects of episodic future thinking on discounting: Personalized age-progressed pictures improve risky long-term health decisions. *Journal of Applied Behavior Analysis*, *49*, 148–169.

Kaplan, B. A., Reed, D. D., & McKerchar, T. L. (2014). Using a visual analogue scale to assess delay, social, and probability discounting of an environmental loss. *The Psychological Record*, *64*, 261–269.

Karsten, A. M., Carr, J. E., & Lepper, T. L. (2011). Description of a practitioner model for identifying preferred stimuli with individuals with autism spectrum disorders. *Behavior Modification*, *35*(4), 347–369. doi:10.1177/0145445511405184

Kasper-Ferguson, S., & Moxley, A. (2002). Developing a writing package with student graphing of fluency. *Education and Treatment of Children*, *25*, 249–267.

Kastak, C. R., Schusterman, R. J., & Kastak, D. (2001). Equivalence classification by California sea lions using class-specific reinforcers. *Journal of the Experimental Analysis of Behavior*, *76*, 131–158.

Kauffman, J. M. (2005). *Characteristics of emotional and behavioral disorders of children and youth* (8th ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Kauffman, J. M. (2011). *Toward a science of education: The battle between rouge and real science*. Verona, WI: Full Court Press.

Kauffman, J. M., & Landrum, T. J. (2013). *Characteristics of emotional and behavioral disorders of children and youth* (10th ed.). Upper Saddle River, NJ: Pearson.

Kauffman, J. M., Cullinan, D., Scranton, T. R., & Wallace, G. (1972). An inexpensive device for programming ratio reinforcement. *The Psychological Record*, *22*, 543–554.

Kaufman, A. S. (2018). *Contemporary intellectual assessment: Theories, tests, and issues*. New York: Guilford Press.

Kaufman, K. F., & O'Leary, K. D. (1972). Reward, cost, and self-evaluation procedures for disruptive adolescents in a psychiatric hospital school. *Journal of Applied Behavior Analysis*, *5*, 293–309.

Kazdin, A. E. (1973). The effects of vicarious reinforcement on attentive behavior in the classroom. *Journal of Applied Behavior Analysis*, *6*, 77–78.

Kazdin, A. E. (1977). *The token economy: A review and evaluation*. New York: Plenum Press.

Kazdin, A. E. (1978). *History of behavior modification*. Austin, TX: PRO-ED.

Kazdin, A. E. (1979). Unobtrusive measures in behavioral assessment. *Journal of Applied Behavior Analysis*, *12*, 713–724.

Kazdin, A. E. (1982). Observer effects: Reactivity of direct observation. *New Directions for Methodology of Social and Behavioral Science*, *14*, 5–19.

Kazdin, A. E. (2001). *Behavior modification in applied settings* (6th ed.). Belmont, CA: Wadsworth/Thomas Learning.

Kazdin, A. E. (2011). *Single case research designs: Methods for clinical and applied settings* (2nd ed.). New York: Oxford University Press.

Kazdin, A. E. (2013). *Behavior modification in applied settings* (7th ed.). Long Grove, IL: Waveland Press.

Kazdin, A. E., & Polster, R. (1973). Intermittent token reinforcement and response maintenance in extinction. *Behavior Therapy*, *4*, 386–391.

Kee, M., Hill, S. M., & Weist, M. D. (1999). School-based behavior management of cursing, hitting, and spitting in a girl with profound retardation. *Education and Treatment of Children*, *22*(2), 171–178.

Keenan, M. (2018). *The shaping game*. Retrieved January 2, 2018, from http://www.imagesforbehaviouranalysts.com/shaping-game.html

Keenan, M., & Dillenburger, K. (2000). Images of behavior analysis: The shaping game and the behavioral stream. *Behavior and Social Issues*, *10*, 19–38.

Keeney, K. M., Fisher, W. W. Adelinis, J. D., & Wilder, D. A. (2000). The effects of response cost in the treatment of aberrant behavior maintained by negative reinforcement. *Journal of Applied Behavior Analysis*, *33*, 255–258.

Kehle, T. J., Bray, M. A., Theodore, L. A., Jenson, W. R., & Clark, E. (2000). A multi-component intervention designed to reduce disruptive classroom behavior. *Psychology in the Schools*, *37*(5), 475–481.

Kehoe, E. J., & Macrae, M. (1998). Classical conditioning. In W. H. O'Donohue (Ed.). *Learning and behavior therapy* (pp. 36–58). Boston: Allyn and Bacon.

Keintz, K. S., Miguel, C. F., Kao, B., & Finn, H. E. (2011). Using conditional discrimination training to produce emergent relations between coins and their values in children with autism. *Journal of Applied Behavior Analysis*, *44*, 909–913.

Kelleher, R. T., & Gollub, L. R. (1962). A review of positive conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, *5*, 543–597.

Kellems, R. O., & Morningstar, M. E. (2012). Using video modeling delivered through iPods to teach vocational tasks to young adults with autism spectrum disorders. *Career Development and Transition for Exceptional Individuals*, *35*(3), 155–167. doi:10.1177/0885728812443082

Keller, F. S. (1941). Light aversion in the white rat. *Psychological Record*, *4*, 235–250.

Keller, F. S. (1968). "Goodbye, Teacher . . ." *Journal of Applied Behavior Analysis*, *1*, 79–89.

Keller, F. S. (1982). *Pedagogue's progress*. Lawrence, KS: TRI Publications.

Keller, F. S. (1990). Burrhus Frederic Skinner (1904–1990) (a thank you). *Journal of Applied Behavior Analysis*, *23*, 404–407.

Keller, F. S., & Schoenfeld, W. M. (1950). *Principles of psychology*. Acton, MA: Copley Publishing Group.

Kelley, H., & Miltenberger, R. G. (2016). Using video feedback to improve horseback-riding skills. *Journal of Applied Behavior Analysis*, *49*, 138–147. doi:10.1002/jaba.272

Kelley, M. E., LaRue, R. H., Roane, H. S., & Gadaire, D. M. (2011). Indirect behavior assessments: Interviews and rating scales. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 182–190). New York: Guilford Press.

Kelley, M. E., Piazza, C. C., Fisher, W. W., & Oberdorff, A. J. (2003). Acquisition of cup drinking using previously refused foods as positive and negative reinforcement. *Journal of Applied Behavior Analysis*, *36*, 89–93.

Kelley, M. E., Shillingsburg, M. A., & Bowen, C. N. (2016). Stability of daily preference across multiple individuals. *Journal of Applied Behavior Analysis*, *49*, 394–398. doi:10.1002/jaba.288

Kelley, M. L., Jarvie, G. J., Middlebrook, J. L., McNeer, M. F., & Drabman, R. S. (1984). Decreasing burned children's pain behavior: Impacting the trauma of hydrotherapy. *Journal of Applied Behavior Analysis*, *17*, 147–158.

Kellum, K. K., Carr, J. E., & Dozier, C. L. (2001). Response-card instruction and student learning in a college classroom. *Teaching of Psychology*, *28*(2), 101–104.

Kelly, A., & Tincani, M. (2013). Collaborative training and practice among applied behavior analysts who support individuals with autism spectrum disorder. *Education and Training in Autism and Developmental Disabilities*, *48*, 120–131.

Kelly, M. A., Roscoe, E. M., Hanley, G. P., & Schlichenmeyer, K. (2014). Evaluation of assessment methods for identifying social reinforcers. *Journal of Applied Behavior Analysis*, *47*, 113–135.

Kelly, M. B. (1977). A review of the observational data collection and reliability procedures reported in the *Journal of Applied Behavior Analysis*. *Journal of Applied Behavior Analysis*, *10*, 97–101.

Kelshaw-Levering, K., Sterling-Turner, H. E., Henry, J. R., & Skinner, C. H. (2000). Randomized interdependent group contingencies: Group reinforcement with a twist. *Journal of School Psychology*, *37*, 523–534.

Kemp, D. C., & Carr, E. G. (1995). Reduction of severe problem behavior in community employment using a hypothesis-driven multicomponent intervention approach. *Journal of the Association for Persons with Severe Handicaps*, *20*, 229–247.

Kennedy, C. H. (1994). Automatic reinforcement: Oxymoron or hypothetical construct. *Journal of Behavioral Education*, *4*(4), 387–395.

Kennedy, C. H. (2005). *Single-case designs for educational research*. Boston: Allyn and Bacon.

Kennedy, C. H., & Sousa, G. (1995). Functional analysis and treatment of eye poking. *Journal of Applied Behavior Analysis*, *28*, 27–37.

Kennedy, C. H., Itkonen, T, & Lindquist, K. (1994). Nodality effects during equivalence class formation: An extension to sight-word reading and concept development. *Journal of Applied Behavior Analysis*, *27*, 673–683.

Kennedy, C. H., Meyer, K. A., Knowles, T., & Shukla, S. (2000). Analyzing the multiple functions of stereotypical behavior for students with autism: Implications for assessment and treatment. *Journal of Applied Behavior Analysis*, *33*, 559–571.

Kern, L., Dunlap, G., Clarke, S., & Childs, K. E. (1995). Student-assisted functional assessment interview. *Diagnostique*, *19*, 29–39.

Kern, L., Koegel, R., & Dunlap, G. (1984). The influence of vigorous versus mild exercise on autistic stereotyped behaviors. *Journal of Autism and Developmental Disorders*, *14*, 57–67.

Kern, L., Mantegna, M.E., Vorndran, C.E., Bailin, D., & Hilt, A. (2001). Choice of task squence to reduce problem behaviors. *Journal of Positive Behavior Interventions*, *3*, 3–10.

Kerr, M. M., & Nelson, C. M. (2002). *Strategies for addressing behavior problems in the classroom* (4th ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Kestner, K. M., & Peterson, S. M. (2017). A review of resurgence literature with human participants. *Behavior Analysis: Research and Practice*, *17*, 1–17. doi.org/10.1037/bar0000039

Kettering, T. L., Neef, N. A., Kelley, M. E. & Heward, W. L. (2018). A comparison of procedures for unpairing conditioned reflexive motivating operations. *Journal of the Experimental Analysis of Behavior*, *109*, 422–432.

Killeen, P. (2018). The futures of experimental analysis of behavior. *Behavior Analysis: Research and Practice*, *18*(2), 124–133.

Killu, K., Sainato, D. M., Davis, C. A., Ospelt, H., & Paul, J. N. (1998). Effects of high-probability request sequences on preschoolers' compliance and disruptive behavior. *Journal of Behavioral Education*, *8*, 347–368.

Kimball, J. K., & Heward, W. L. (1993). A synthesis of contemplation, prediction, and control. *American Psychologist*, *48*, 587–588.

King, J. E., & Hayes, L. J. (2016). The role of discriminative stimuli on response resurgence. *The Psychological Record*, *66*, 325–335. doi.10.1007/s40732-016-0175-2

Kinloch, J. M., Foster, M. T., & McEwan, J. S. A. (2009). Extinction-induced variability in human behavior. *The Psychological Record*, *59*, 347–370.

Kirby, K. C., & Bickel, W. K. (1988). Toward an explicit analysis of generalization: A stimulus control interpretation. *The Behavior Analyst*, *11*, 115–129.

Kirby, K. C., Fowler, S. A., & Baer, D. M. (1991). Reactivity in self-recording: Obtrusiveness of recording procedure and peer comments. *Journal of Applied Behavior Analysis*, *24*, 487–498.

Kirby, K. N., Petry, N. M., & Bickel, W. K. (1999). Heroin addicts have higher discounting rates for delayed rewards than non-drug-using controls. *Journal of Experimental Psychology: General*, *128*, 78–87.

Kirigan, K. A., Braukmann, C. J., Atwater, J. D., and Wolf, M. M. (1982). An evaluation of teaching family (Achievement Place) group homes for juvenile offenders. *Journal of Applied Behavior Analysis*, *15*, 1–16.

Kisamore, A. N., Carr, J. E., & LeBlanc, L. A. (2011). Training preschool children to use visual imagining as a problem-solving strategy for complex categorization tasks. *Journal of Applied Behavior Analysis*, *44*, 255–278. doi.org/10.1901/jaba.2011.44-255

Kisamore, A. N., Karsten, A. M., & Mann, C. C. (2016). Teaching multiply controlled intraverbals to children and adolescents with autism spectrum disorders. *Journal of Applied Behavior Analysis*, *49*, 826–847.

Kladopoulos, C. N., & McComas, J. J. (2001). The effects of form training on foul-shooting performance in members of a women's college basketball team. *Journal of Applied Behavior Analysis*, *34*, 329–332. doi:10.1901/jaba.2001.34-329

Klatt, K. P., Sherman, J. A., & Sheldon, J. B. (2000). Effects of deprivation on engagement in preferred activities by persons with developmental disabilities. *Journal of Applied Behavior Analysis*, *33*, 495–506.

Klein, L. A., Houlihan, D., Vincent, J. L., & Panahon, C. J. (2017). Best practices in utilizing the changing criterion design. *Behavior Analysis in Practice*, *1*(1), 52–61.

Kleinman, K. E., & Saigh, P. A. (2011). The effects of the Good Behavior Game on the conduct of regular education New York City high school students. *Behavior Modification*, *35*(1), 95–105. doi:10.1177/0145445510392213

Kleinmann, A. E., Luiselli, J. K., DiGennaro, F. D., Pace, G. M., Langone, S. R., & Chochran, C. (2009). Systems-level assessment of interobserver agreement (IOA) for implementation of protective holding (therapeutic restraint) in a behavioral healthcare setting. *Journal of Developmental and Physical Disabilities*, *21*, 473–483.

Klimas, A., & McLaughlin, T. F. (2007). The effects of a token economy system to improve social and academic behavior with a rural primary aged child with disabilities. *International Journal of Special Education*, *22*(3), 72–77. Retrieved from http://eric.ed.gov/?id=EJ814513

Kneedler, R. D., & Hallahan, D. P. (1981). Self-monitoring of on-task behavior with learning disabled children: Current studies and directions. *Exceptional Education Quarterly*, *2*(3), 73–82.

Knight, M. G., Ross, D. E., Taylor, R. L., & Ramasamy, R. (2003). Constant time delay and interpersal of known items to teach sight words to students with mental retardation and learning disabilities. *Education and Training in Developmental Disabilities*, *38*, 179–191.

Knittel, D. (2012). Case study: A professional guitarist on comeback road. In R. W. Malott (Ed.), *I'll stop procrastinating when I get around to it: Plus other cool ways to succeed in school and life using behavior analysis to get your act together* (pp. 6–82, 6–83). Kalamazoo, MI: Behaviordelia.

Knudson, P. J., Miltenberger, R. G., Bosch, A., Gross, A., Brower-Breitwieser, & Tarasendko, M. (2009). Fire safety skills training for individuals with severe and profound mental retardation. *Journal of Developmental and Physical Disabilities*, *21*, 523–535. doi:10.1007/s10882-009-9161-9

Kobari-Wright, V. V., & Miguel, C. F. (2014). The effects of listener training on the emergence of categorization and speaker behavior in children with autism. *Journal of Applied Behavior Analysis*, *47*, 431–436.

Kodak, T., Campbell, V., Bergmann, S., LeBlanc, B., Kurtz-Nelson, E., Cariveau, T., & Mahon, J. (2016). Examination of efficacious, efficient, and socially valid error-correction procedures to teach sight words and prepositions to children with autism spectrum disorder. *Journal of Applied Behavior Analysis*, *49*, 532–547.

Kodak, T., Grow, L., & Northrup, J. (2004). Functional analysis and treatment of elopement for a child with attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis*, *37*, 229–232.

Kodak, T., Miltenberger, R. G., & Romaniuk, C. (2003). The effects of differential reinforcement of other behavior and noncontingent escape on compliance. *Journal of Applied Behavior Analysis*, *36*, 379–382. doi.org/10.1901/jaba.2003.36-379

Koegel, L. K., Carter, C. M., & Koegel, R. L. (2003). Teaching children with autism self-initiations as a pivotal response. *Topics in Language Disorders*, *23*(2), 134–145.

Koegel, L. K., Koegel, R .L., Green-Hopkins, I., & Barnes, C. C. (2010). Brief report: Question- asking and collateral language acquisition in children with autism. *Journal of Autism and Developmental Disorders*, *40*(4), 509–515.

Koegel, L. K., Koegel, R. L., Hurley, C., & Frea, W. D. (1992). Improving social skills and disruptive behavior in children with autism through self-management. *Journal of Applied Behavior Analysis*, *25*, 341–353.

Koegel, L. K., Park, M. N., & Koegel, R. L. (2014). Using self-management to improve the reciprocal social conversation of children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *44*(5), 1055–1063.

Koegel, R. L., & Koegel, L. K. (1988). Generalized responsivity and pivotal behaviors. In R. H. Horner, G. Dunlap, & R. L. Koegel (Eds.), *Generalization and maintenance: Life-style changes in applied settings* (pp. 41–66). Baltimore: Paul H. Brookes.

Koegel, R. L., & Koegel, L. K. (2018). *Pivotal response treatment for autism spectrum disorders* (2nd ed.). Baltimore: Brookes.

Koegel, R. L., & Rincover, A. (1977). Research on the differences between generalization and maintenance in extra-therapy responding. *Journal of Applied Behavior Analysis*, *10*, 1–12.

Koegel, R. L., & Williams, J. A. (1980). Direct versus indirect response-reinforcer relationships in teaching autistic children. *Journal of Abnormal Child Psychology*, *8*, 537–547.

Koegel, R. L., Bradshaw, J. L., Ashbaugh, K., & Koegel, L. K. (2014). Improving question-asking initiations in young children with autism using pivotal response treatment. *Journal of Autism and Developmental Disorders*, *44*(4), 816–827.

Koegel, R. L., Koegel, L. K., & Schreibman, L. (1991). Assessing and training parents in teaching pivotal behaviors. In R. J. Prinz (Ed.), *Advances in behavioral assessment of children and families* (Vol. 5, pp. 65–82). London, UK: Jessica Kingsley.

Koegel, R. L., Openden, D., & Koegel, L. K. (2004). A systematic desensitization paradigm to treat hypersensitivity to auditory stimuli in children with

autism in family contexts. *Research and Practice for Persons with Severe Disabilities*, *29*, 122–134.

Koehler, L. J., Iwata, B. A., Roscoe, E. M., Rolider, N. U., & O'Steen, L. E. (2005). Effects of stimulus variation on the reinforcing capability of nonpreferred stimuli. *Journal of Applied Behavior Analysis*, *38*, 469–484.

Koehler-Platten, K., Grow, L. L., Schulze, K. A., & Bertone, T. (2013). Using a lag reinforcement schedule to increase phonemic variability in children with autism spectrum disorders. *The Analysis of Verbal Behavior*, *29*, 71–83.

Kohlenberg, B. S., Hayes, S. C., & Hayes, L. J. (1991). The transfer of contextual control over equivalence classes through equivalence classes: A possible model of social stereotyping. *Journal of the Experimental Analysis of Behavior*, *56*, 505–518.

Kohler, F. W., Strain, P. S., Maretsky, S., & Decesare, L. (1990). Promoting positive and supportive interactions between preschoolers: An analysis of group-oriented contingencies. *Journal of Early Intervention*, *14*(4), 327–341.

Kohler, K. T., & Malott, R. W. (2014). Matrix training and verbal generativity in children with autism. *The Analysis of Verbal Behavior*, *30*, 170–177. doi:10.1007/s40616-014-0016-9

Kohn, A. (1997). Students don't "work"—they learn. *Education Week*, September 3, 1997. Retrieved January 4, 2017, from www.alfiekohn.org/article/students-dont-work-learn/.

Kohn, C. S. (2006). Conceptualization and treatment of kleptomania behaviors using cognitive and behavioral strategies. *International Journal of Behavioral Consultation and Therapy*, *2*(4), 553–559.

Komaki, J. L. (1998). When performance improvement is the goal: A new set of criteria for criteria. *Journal of Applied Behavior Analysis*, *31*, 263–280.

Konarski, E. A., Jr., Crowell, C. R., & Duggan, L. M. (1985). The use of response deprivation to increase the academic performance of EMR students. *Applied Research in Mental Retardation*, *6*, 15–31.

Konarski, E. A., Jr., Crowell, C. R., Johnson, M. R., & Whitman T. L. (1982). Response deprivation, reinforcement, and instrumental academic performance in an EMR classroom. *Behavior Therapy*, *13*, 94–102.

Konarski, E. A., Jr., Johnson, M. R., Crowell, C. R., & Whitman T. L. (1980). Response deprivation and reinforcement in applied settings: A preliminary report. *Journal of Applied Behavior Analysis*, *13*, 595–609.

Koocher, G. P., & Gill, E. (2016). Pet me, sniff me, squeeze me. In R. M. Foxx & J. A. Mulick (Eds.), *Controversial therapies for autism and intellectual disabilities: Fads, fashion, and science in professional practice* (2nd ed.) (pp. 136–153). New York: Routledge.

Koocher, G. P., & Keith-Spiegel, P. (1998). *Ethics in psychology: Professional standards and cases*. New York, Oxford: Oxford University Press.

Kortick, S. A., & O'Brien, R. M. (1996). The world series of quality control: A case study in the package delivery industry. *Journal of Organizational Behavior Management*, *16*(2), 77–93.

Kosiewicz, M. M., Hallahan, D. P., Lloyd, J. W., & Graves, A. W. (1982). Effects of self-instruction and self-correction procedures on handwriting performance. *Learning Disability Quarterly*, *5*(1), 71–78.

Kostewicz, D. E., King, S. A., Datchuk, S. M., Brennan, K. M., & Casey, S. D. (2016). Data collection and measurement assessment in behavioral research: 1958–2013. *Behavior Analysis: Research and Practice*, *16*(1), 19–33. doi.org/10.1037/bar0000031

Kostewicz, D. E., Kubina, R. M., Jr., & Cooper, J. O. (2000). Managing aggressive thoughts and feelings with daily counts of non-aggressive thoughts and feelings: A self-experiment. *Journal of Behavior Therapy and Experimental Psychiatry*, *31*, 177–187.

Kounin, J. (1970). *Discipline and group management in classrooms*. New York: Holt, Rinehart & Winston.

Kranak, M. C., Alber-Morgan, S. R., & Sawyer, M. R. (2017). A parametric analysis of specific praise rates on the on-task behavior of elementary students with autism. *Education and Treatment of Autism and Developmental Disabilities*, *52*(4), 453–464.

Krantz, P. J., & McClannahan, L. E. (1998). Social interaction skills for children with autism: A script-fading procedure for beginning readers. *Journal of Applied Behavior Analysis*, *31*, 191–202.

Krasner, L. A., & Ullmann, L. P. (Eds.). (1965). *Research in behavior modification: New developments and implications*. New York: Holt, Rinehart & Winston.

Kratochwill, T. R., & Levin, J. R. (2014). Meta- and statistical analysis of single-case intervention research data: Quantitative gifts and a wish list. *Journal of School Psychology*, *52*, 123–127.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shaddish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education*, *34*, 26–38.

Krentz, H., Miltenberger, R., & Valbuena, D. (2016). Using token reinforcement to increase walking for adults with intellectual disabilities. *Journal of Applied Behavior Analysis*, *49*, 745–750.

Kretlow, A. G., Cooke, N. L., & Wood, C. L. (2012). Using in-service and coaching to increase teachers' accurate use of research-based strategies. *Remedial and Special Education*, *33*(6), 348–361.

Krueger, T. K., Rapp, J. T., Ott, L. M., Lood, E. A., & Novotny, M. A. (2013). Detecting false positives in A-B designs: Potential implications for practitioners. *Behavior Modification*, *37*(5), 615–630. doi:10.1177/0145445512468754

Kubina, R. M., Jr. (2005). The relations among fluency, rate building, and practice: A response to Doughty, Chase, and O'Shields (2004). *The Behavior Analyst*, *28*, 73–76.

Kubina, R. M., & Yurich, K. K. L. (2012). *The precision teaching book*. Lemont, PA: Greatness Achieved.

Kubina, R. M., Haertel, M. W., & Cooper, J. O. (1994). Reducing negative inner behavior of senior citizens: The one-minute counting procedure. *Journal of Precision Teaching*, *11*(2), 28–35.

Kubina, R. M., Kostewicz, D. E., Brennan, K. M., & King, S. A. (2015). A critical review of line graphs in behavior analytic journals. *Educational Psychology Review*, *29*, 583–598.

Kubina, R. M., Kostewicz, D. E., & Datchuk, S. M. D. (2008). An initial survey of fractional graph and table area in behavioral journals. *The Behavior Analyst*, *31*, 61–66.

Kuhl, S., Rudrud, E. H., Witts, B. N., & Schulze, K. A. (2015). Classroom-based interdependent group contingencies increase children's physical activity. *Journal of Applied Behavior Analysis*, *48*, 602–612.

Kuhn, S. A. C., Lerman, D. C., Vorndran, C. M., & Addison, L. (2006). Analysis of factors that affect responding in a two-response chain in children with developmental disabilities. *Journal of Applied Behavior Analysis*, *39*, 263–280.

Kuoch, H., & Mirenda, P. (2003). Social story interventions for young children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, *18*, 219–227.

Kurti, A. N., & Dallery, J. (2013). Internet-based contingency management increases walking in sedentary adults. *Journal of Applied Behavior Analysis*, *46*, 568–581.

Kurtz, P. F., & Lind, M. A. (2013). Behavioral approaches to treatment of intellectual and developmental disabilities. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. 279–299). Washington, DC: American Psychological Association.

La Greca, A. M., & Schuman, W. B. (1995). Adherence to prescribed medical regimens. In M. C. Roberts (Ed.), *Handbook of pediatric psychology* (2nd ed.) (pp. 55–83). New York: Guilford Press.

La Londe, K. B., Mahoney, A., Edwards, T. L., Cox, C., Weetjens, B. J., Durgin, A., & Poling, A. (2010). Teaching pouched rats to find people. *Journal of Applied Behavior Analysis*, *48*, 1–10. doi:10.1002/jaba.181

LaFrance, D. L., & Tarbox, J. (in press). The importance of multiple exemplar instruction in the establishment of novel verbal behavior. *Journal of Applied Behavior Analysis*.

LaFrance, D. L., Tarbox, J., & Holth, P. (2019). *On terms: Multiple exemplar instruction and multiple exemplar training*. Manuscript in progress.

Lagorio, C. H., & Madden, G. J. (2005). Delay discounting of real and hypothetical rewards: III. Steady-state assessments, forced-choice trials, and all real rewards. *Behavioural Processes*, *69*, 173–187.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Lalli, J. S., Browder, D. M., Mace, F. C., & Brown, D. K. (1993). Teacher use of descriptive analysis data to implement interventions to decrease students' problem behaviors. *Journal of Applied Behavior Analysis*, *26*, 227–238.

Lalli, J. S., Casey, S., & Kates, K. (1995). Reducing escape behavior and increasing task completion with functional communication training, extinction, and response chaining. *Journal of Applied Behavior Analysis*, *28*, 261–268.

Lalli, J. S., Livezey, K., & Kates, D. (1996). Functional analysis and treatment of eye poking with response blocking. *Journal of Applied Behavior Analysis*, *29*, 129–132.

Lalli, J. S., Mace, F. C., Livezey, K., & Kates, K. (1998). Assessment of stimulus generalization gradients in the treatment of self-injurious behavior. *Journal of Applied Behavior Analysis*, *31*, 479–483.

Lam, A. L., Cole, C. L., Shapiro, E. S., & Bambara, L. M. (1994). Relative effects of self-monitoring on-task behavior, academic accuracy, and disruptive behavior in students with behavior disorders. *School Psychology Review*, *23*, 44–59.

Lamb, R. J., Kirby, K. C., Morral, A. R., Galbicka, G., & Iguchi, M. Y. (2010). Shaping smoking cessation in hard-to-treat smokers. *Journal of Consulting and Clinical Psychology*, *78*, 62–71. doi:10.1037/a0018323.

Lamb, R. J., Morral, A. R., Kirby, K. C., Iguchi, M. Y., & Galbicka, G. (2004). Shaping smoking cessation using percentile schedules. *Drug and Alcohol Dependence*, *76*, 247–259.

Lambert, J. M., Bloom, S. E., & Irvin, J. (2012). Trial-based functional analysis and functional communication training in an early childhood setting. *Journal of Applied Behavior Analysis*, *45*, 579–584. doi:10.1901/jaba.2012.45-579

Lambert, J. M., Bloom, S. E., Samaha, A. L., Dayton, E., & Rodewald, A. (2015). Serial alternative response training as intervention for target response resurgence. *Journal of Applied Behavior Analysis*, *48*, 765–780.

Lambert, J. M., Copeland, B. A., Karp, E. L., Finley, C. I., Houchins-Juarez, N. J. & Ledford, J. R. (2016). Chaining functional basketball sequences (with embedded conditional discriminations) in an adolescent with autism. *Behavior Analysis in Practice*, *9*, 199–210. doi:10.1007/s40617-016-0125-0

Lambert, M. C., Cartledge, G., Lo, Y., & Heward, W. L. (2006). Effects of response cards on disruptive behavior and academic responding by fourth-grade urban students. *Journal of Positive Behavioral Interventions*, *8*, 88–99.

Landa, R. K., Hansen, B., & Shillingsburg, B. M. (2017). Teaching mands for information using 'when' to children with autism. *Journal of Applied Behavior Analysis*, *50*, 538–551.

Lane, S. D., & Critchfield, T. S. (1998). Classification of vowels and consonants by individuals with moderate mental retardation: Development of arbitrary relations via match-to-sample training with compound stimuli. *Journal of Applied Behavior Analysis*, *31*, 21–41.

Lang, R., O'Reilly, M., Sigafoos, J., Lancioni, G. E., Machalicek, W., Rispoli, M., & White, P. (2009). Enhancing the effectiveness of a play intervention by abolishing the reinforcing value of stereotypy: A pilot study. *Journal of Applied Behavior Analysis*, *42*, 889–894.

Lannie, A. L., & McCurdy, B. L. (2007). Preventing disruptive behavior in the urban classroom: Effects of the Good Behavior Game on student and teacher behavior. *Education and Treatment of Children*, *30*(1), 85–98.

Lanovaz, M. J., Rapp, J. T., Long, E. S., Richling, S. M., & Carroll, R. A. (2014). Preliminary effects of conditioned establishing operations on stereotypy. *The Psychological Record*, *64*, 209–216.

Laraway, S., Snycerski, S., Michael, J., & Poling, A. (2001). The abative effect: A new term to describe the action of antecedents that reduce operant responding. *The Analysis of Verbal Behavior*, *18*, 101–104.

Laraway, S., Snycerski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them: Some further refinements. *Journal of Applied Behavior Analysis*, *36*, 407–414. doi.org/10.1901/jaba.2003.36-407

Laraway, S., Snycerski, S., Olson, R., Becker, B., & Poling, A. (2014). The motivating operations concept: Current status and critical response. *The Psychological Record*, *64*, 601–623.

LaRowe, L. N., Tucker, R. D., & McGuire, J. M. (1980). Lunchroom noise control using feedback and group reinforcement. *Journal of School Psychology*, *18*, 51–57.

Lasiter, P. S. (1979). Influence of contingent responding on schedule-induced activity in human subjects. *Physiology and Behavior*, *22*, 239–243.

Lassman, K. A., Jolivette, K., & Wehby, J. H. (1999). "My teacher said I did good work today!": Using collaborative behavioral contracting. *Teaching Exceptional Children*, *31*(4), 12–18.

Lattal, A. (2013). The five pillars of the experimental analysis of behavior. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. 33–64). Washington, DC: American Psychological Association.

Lattal, A. D., & Clark, R. W. (2007). *A good day's work: Sustaining ethical behavior and business success*. New York: McGraw-Hill.

Lattal, K. A. (1969). Contingency management of toothbrushing behavior in a summer camp for children. *Journal of Applied Behavior Analysis*, *2*, 195–198.

Lattal, K. A. (1995). Contingency and behavior analysis. *The Behavior Analyst*, *18*, 209–224.

Lattal, K. A. (2004). Steps and pips in the history of the cumulative recorder. *Journal of Experimental Analysis of Behavior*, *82*, 329–355. doi:10.1901/jeab.2004.82-329

Lattal, K. A., & Gleeson, S. (1990). Response acquisition with delayed reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, *16*, 27–39.

Lattal, K. A., & Griffin, M. A. (1972). Punishment contrast during free operant responding. *Journal of the Experimental Analysis of Behavior*, *18*, 509–516.

Lattal, K. A., & Lattal, A. D. (2006). And yet . . . : Further comments on distinguishing positive and negative reinforcement. *The Behavior Analyst*, *29*, 129–134.

Lattal, K. A., & Neef, N. A. (1996). Recent reinforcement-schedule research and applied behavior analysis. *Journal of Applied Behavior Analysis*, *29*, 213–230.

Lattal, K. A., & Shahan, T. A. (1997). Differing views of contingencies: How contiguous. *The Behavior Analyst*, *20*, 149–154.

Lattal, K. A., & St. Peter Pipkin, C. (2009). Resurgence of previously reinforced responding: Research and application. *The Behavior Analyst Today*, *10*, 254–266.

Lattal, K. A., & Ziegler, D. R. (1982). Briefly delayed reinforcement: An interresponse time analysis. *Journal of the Experimental Analysis of Behavior*, *37*, 407–416.

Lattal, K. A., St. Peter Pipkin, C., & Escobar, R. (2013). Operant extinction. In G. J. Madden (Ed.), *APA handbook of behavior analysis, Vol. 2: Translating principles into practice* (pp. 77–107). Washington, DC: American Psychological Association.

Lattal, K. M. (2013). Pavlovian conditioning. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. 283–306). Washington, DC: American Psychological Association.

Lavie, T., & Sturmey, P. (2002). Training staff to conduct a paired-stimulus preference assessment. *Journal of Applied Behavior Analysis*, *35*, 209–211. doi:10.1901/jaba.2002.35-209

LaVigna, G. W., & Donnellen, A. M. (1986). *Alternatives to punishment: Solving behavior problems with nonaversive strategies*. New York: Irvington.

Layng, T. V. J. (2017). Private emotions as contingency descriptors: Emotions, emotional behavior, and their evolution. *European Journal of Behavior Analysis*, *18*, 168–179. doi.org/10.1080/15021149.2017.1304875

Layng, T. V. J. (2017). *Understanding concepts: Implications for science teaching*. Generategy, LLC.

Layng, T. V. J., & Andronis, P. T. (1984). Toward a functional analysis of delusional speech and hallucinatory behavior. *The Behavior Analyst*, *7*, 139–156.

Layng, T. V. J., Twyman, J. S., & Stikeleather, G. (2004). Engineering discovery learning: The contingency adduction of some precursors of textual responding in a beginning reading program. *The Analysis of Verbal Behavior*, *20*, 99–109.

Layng, T. V. J., Twyman, J. S., & Stikeleather, G. (2004). Selected for success: How Headsprout Reading Basics™ teaches children to read. In D. J. Moran and R. W. Malott (Eds.) *Evidence based education methods* (pp. 171–197). St. Louis, MO: Elsevier/Academic Press.

Lazar, R. M., Davis-Lang, D., & Sanchez, L. (1984). The formation of visual stimulus equivalences in children. *Journal of the Experimental Analysis of Behavior*, *41*, 251–266.

Le, D. D., & Smith, R. G. (2002). Functional analysis of self-injury with and without protective equipment. *Journal of Developmental and Physical Disabilities*, *14*, 277–290.

Leader, G., & Barnes-Holmes, D. (2001a). Establishing fraction-decimal equivalence using a respondent-type training procedure. *The Psychological Record*, *51*, 151–165.

Leader, G., & Barnes-Holmes, D. (2001b). Matching-to-sample and respondent-type training as methods for producing equivalence relations: Isolating the critical variable. *The Psychological Record*, *51*, 429–444.

Leaf, J. B., Leaf, R. B., Alcalay, A., Leaf, J. A., Ravid, D., Dale, S., . . . Oppenheim-Leaf, M. L. (2015). The utility of formal preference assessments for individuals diagnosed with autism spectrum disorder. *Education and Training in Autism and Developmental Disabilities*, *50*, 199–212.

Leaf, J. B., Leaf, R., Leaf , J. A., Alcalay, A., Ravid, D. Dale, S., . . . Oppenheim-Leaf, M. L. (2018). Comparing paired-stimulus preference assessments with in-the-moment reinforcer analysis on skill acquisition: A preliminary investigation. *Focus on Autism and Other Developmental Disabilities*, *33*, 14–24. doi:10.1177/1088357616645329

Leaf, J. B., Oppenheim-Leaf, M. L., Leaf, R., Courtermanche, A. B., Taubman, M., McEachin, J., . . . Sherman, J. A. (2012). Observational effects on the preferences of children with autism. *Journal of Applied Behavior Analysis, 45*, 473–483.

Leaf, R., & McEachin, J. (1999). *A work in progress*. New York: DRL Books.

Leão, M. F. F. C., Laurenti, C., & Haydu, V. B. (2016). Darwinism, radical behaviorism, and the role of variation in Skinnerian explaining behavior. *Behavior Analysis: Research and Practice, 16*(1), 1–11. doi.org/10.1037/bar0000025

Lebbon, A., Austin, J., Van Houten, R., & Malenfant, J. E. (2007). Evaluating the effects of traffic on driver stopping and turn signal use at a stop sign. *Journal of Organizational Behavior Management, 27*(2), 27–35. doi:10.1300/J075v27n02_03

LeBlanc, L. A., & Dillon, C. M. (2009). Establishing mand and tact repertoires. In R. A. Rehfeldt & Y. Barnes-Holmes (Eds.), *Derived relational responding applications for learners with autism and other developmental disabilities* (pp. 79–108). Oakland, CA: New Harbinger Publications, Inc.

LeBlanc, L. A., & Luiselli, J. K. (2016). Refining supervisory practices in the field of behavior analysis: Introduction to the special section on supervision. *Behavior Analysis in Practice, 9*, 271–273. doi 10.1007/s40617-016-0156-6

LeBlanc, L. A., Cherup, S. M., Feliciano, L., & Sidener, T. M. (2006). Using choice-making opportunities to increase activity engagement in individuals with dementia. *American Journal of Alzheimer's Disease and Other Dementias, 21*, 318–325.

LeBlanc, L. A., Coates, A. M., Daneshvar, S., Charlop-Christy, M. H., Morris, C., & Lancaster, B. M. (2003). Using video modeling and reinforcement to teach perspective-taking skills to children with autism. *Journal of Applied Behavior Analysis, 36*, 253–257.

LeBlanc, L. A., Heinicke, M. R., & Baker, J. C. (2012). Expanding the consumer base for behavior analytic services: Meeting the needs of consumers in the 21st century. *Behavior Analysis in Practice, 5*, 4–14.

LeBlanc, L. A., Miguel, C. F., Cummings, A. R., Goldsmith, T. R., & Carr, J. E. (2003). The effects of three stimulus-equivalence testing conditions on emergent US geography relations of children diagnosed with autism. *Behavioral Interventions, 18*, 279–289.

LeBlanc, L. A., Raetz, P. B., & Feliciano, L. (2011). Behavioral gerontology. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 472–488). New York: Guilford Press.

LeBlanc, L. A., Raetz, P. B., Sellers, T. P., & Carr, J. E. (2016). A proposed model for selecting measurement procedures for the assessment and treatment of problem behavior. *Behavior Analysis in Practice, 9*(1), 77–83.

Lechago, S. A., & Low, A. (2015). A review of the mand-for-information training research literature. *International Journal of Behavior Analysis & Autism Spectrum Disorders, 1*, 35–54.

Lechago, S. A., Carr, J. E., Grow, L. L., Love, J. R., & Almason, S. M. (2010). Mands for information generalize across establishing operations. *Journal of Applied Behavior Analysis, 43*, 381–395. doi.org/10.1901/jaba.2010.43-381

Lechago, S. A., Carr, J. E., Kisamore, A. N., & Grow, L. L. (2015). The effects of multiple exemplar instruction on the relation between listener and intraverbal categorization repertoires. *The Analysis of Verbal Behavior, 31*, 76–95. doi.org/10.1007/s40616-015-0027-1

Ledford, J. R., & Gast, D. L. (2014). Measuring procedural fidelity in behavioural research. *Neuropsychological Rehabilitation, 24*(3-4), 332–348. http://dx.doi.org/10.1080/09602011.2013.861352

Ledford, J. R., & Wolery, M. (2013). Procedural fidelity: An analysis of measurement and reporting practices. *Journal of Early Intervention, 35*(2), 173–193.

Ledford, J. R., & Wolery, M. (2013). Effects of plotting a second observer's data on A-B-A-B graphs when observer disagreement is present. *Journal of Behavioral Education, 22*, 312–324. doi:10.1007/s10864-013-9178-0

Ledford, J. R., Artman, K., Wolery, M., & Wehby, J. (2012). The effects of graphing a second observer's data on judgments of functional relations for A-B-A-B graphs. *Journal of Behavioral Education, 21*, 350–364.

Lee, D. L., Belfiore, P. J., Scheeler, M. C., Hua, Y., & Smith, R. (2004). Behavioral momentum in academics: Using embedded high-p sequence to increase academic productivity. *Psychology in the Schools, 41*, 789–801.

Lee, G. P., Miguel, C. F., Darcey, E. K., & Jennings, A. M. (2015). A further evaluation of the effects of listener training on the emergence of speaker behavior and categorization in children with autism. *Research in Autism Spectrum Disorders, 19*, 72–81. doi:10.1016/j.rasd.2015.04.007

Lee, K., Penrod, B., & Price, J. N. (2017). A comparison of cost and reward procedures with interdependent group contingencies. *Behavior Modification, 41*(1), 21–44.

Lee, M. S. H., Yu, C. T., Martin, T. L., & Martin, G. L. (2010). On the relation between reinforcer efficacy and preference. *Journal of Applied Behavior Analysis, 43*, 95–100.

Leeming, E. M., Hansen, D., Alavosius, M., & Reimer, D. (2013). Sustainability in the field: Lake Tahoe hospitality and environmental protection. *Behavior and Social Issues, 22*, 21–39.

Legge, D. B., DeBar, R. M., & Alber-Morgan, S. R. (2010). The effects of self-monitoring with a MotivAider [R] on the on-task behavior of fifth and sixth graders with autism and other disabilities. *Journal of Behavior Assessment and Intervention in Children, 1*(1), 43–52.

Leigland, S. (1984). On "Setting events" and related concepts. *The Behavior Analyst, 7*, 41–45.

Leigland, S. (2010). Functions of research in radical behaviorism for the further development of behavior analysis, *The Behavior Analyst, 33*(2), 207–222.

Leitenberg, H. (1973). The use of single-case methodology in psychotherapy research. *Journal of Abnormal Psychology, 82*, 87–101.

Lennox, D. B., Miltenberger, R. G., & Donnelly, D. R. (1987). Response interruption and DRL for the reduction of rapid eating. *Journal of Applied Behavior Analysis, 20*, 279–284.

Lenz, M., Singh, N. N., & Hewett, A. E. (1991). Overcorrection as an academic remediation procedure: A review and reappraisal. *Behavior Modification, 15*, 64–73.

Lerman, D., & Toole, L. M. (2011). Developing function-based punishment procedures for problem behavior. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 348–369). New York: Guilford Press.

Lerman, D., Hovanetz, A., Strobel, M., & Tetreault, A. (2009). Accuracy of teacher-collected descriptive analysis data: A comparison of narrative and structured recording formats. *Journal of Behavioral Education, 18*, 157–172. Retrieved from http://www.jstor.org.libproxy.library.wmich.edu/stable/41824458

Lerman, D. C. (2003). From the laboratory to community application: Translational research in behavior analysis. *Journal of Applied Behavior Analysis, 36*, 415–419.

Lerman, D. C., & Iwata, B. A. (1993). Descriptive and experimental analysis of variables maintaining self-injurious behavior. *Journal of Applied Behavior Analysis, 26*, 293–319.

Lerman, D. C., & Iwata, B. A. (1995). Prevalence of the extinction burst and its attenuation during treatment. *Journal of Applied Behavior Analysis, 28*, 93–94.

Lerman, D. C., & Iwata, B. A. (1996). A methodology for distinguishing between extinction and punishment effects associated with response blocking. *Journal of Applied Behavior Analysis, 29*, 231–234.

Lerman, D. C., & Iwata, B. A. (1996a). Developing a technology for the use of operant extinction in clinical settings: An examination of basic and applied research. *Journal of Applied Behavior Analysis, 29*, 345–382.

Lerman, D. C., & Vorndran, C. M. (2002). On the status of knowledge for using punishment: Implications for treating behavior disorders. *Journal of Applied Behavior Analysis, 35*, 431–464.

Lerman, D. C., Dittlinger, L. H., Fentress, G., & Lanagan, T. (2011). A comparison of methods for collecting data on performance during discrete trial teaching. *Behavior Analysis in Practice, 4*, 53–62.

Lerman, D. C., Iwata, B. A., & Hanley, G. P. (2013). Applied behavior analysis. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. 81–104). Washington, DC: American Psychological Association.

Lerman, D. C., Iwata, B. A., Shore, B. A., & DeLeon, I. G. (1997). Effects of intermittent punishment on self-injurious behavior: An evaluation of schedule thinning. *Journal of Applied Behavior Analysis, 30*, 187–201. doi:10.1901/jaba.1997.30-187

Lerman, D. C., Iwata, B. A., Shore, B. A., & Kahng, S. (1996). Responding maintained by intermittent reinforcement: Implications for the use of extinction with problem behavior in clinical settings. *Journal of Applied Behavior Analysis*, 29, 153–171. doi:10.1901/jaba.1996.29-153

Lerman, D. C., Iwata, B. A., Smith, R. G., Zarcone, J. R., & Vollmer, T. R. (1994). Transfer of behavioral function as a contributing factor in treatment relapse. *Journal of Applied Behavior Analysis, 27*, 357–370.

Lerman, D. C., Iwata, B. A., & Wallace, M. D. (1999). Side effects of extinction: Prevalence of bursting and aggression during the treatment of self-injurious behavior. *Journal of Applied Behavior Analysis*, *32*, 1–8.

Lerman, D. C., Iwata, B. A., Zarcone, J. R., & Ringdahl, J. (1994). Assessment of stereotypic and self-injurious behavior as adjunctive responses. *Journal of Applied Behavior Analysis*, *27*, 715–728.

Lerman, D. C., Kelley, M. E., Vorndran, C. M., Kuhn, S. A. C., & LaRue, R. H., Jr. (2002). Reinforcement magnitude and responding during treatment with differential reinforcement. *Journal of Applied Behavior Analysis*, *35*, 29–48.

Lerman, D. C., Kelley, M. E., Vorndran, C. M., & Van Camp, C. M. (2003). Collateral effects of response blocking during the treatment of stereotypic behavior. *Journal of Applied Behavior Analysis*, *36*, 119–123. doi.org/10.1901/jaba.2003.36-119

Levin, D., Volkert, V. M., & Piazza, C. C. (2014). A multicomponent treatment package to reduce packing in children with feeding and autism spectrum disorders. *Behavior Modification*, *38*, 940–963. doi:10.1177/0145445514550683

Levy, I. M., Pryor, K. W., & McKeon, T. R. (2016). Is teaching simple surgical skills using an operant learning program more effective than teaching by demonstration? *Clinical Orthopaedics and Related Research*, *474*(4), 945–955. doi:10.1007/s11999-015-4555-8.

Lewis, T. J., Powers, L. J., Kelk, M. J., & Newcomer, L. L. (2002). Reducing problem behaviors on the playground: An intervention of the application of school-wide positive behavior supports. *Psychology in the Schools*, *39*(2), 181–190.

Lewis, T., Scott, T., & Sugai, G. (1994). The problem behavior questionnaire: A teacher-based instrument to develop functional hypotheses of problem behavior in general education classrooms. *Diagnostique*, *19*, 103–115.

Lewon, M., & Hayes, L. J. (2015). The effect of the magnitude of the food deprivation motivating operation on free operant preference in mice. *Behavioral Processes*, *115*, 135–142.

Li, A., Mahoney, A., & Poling, A. (2018). Basic research in behavior analysis. *Behavior Analysis: Research and Practice*, *18*(2), 117–118.

Li, A., Wallace, L., Ehrhardt, K. E., & Poling, A. (2017). Reporting participant characteristics in intervention articles published in five behavior-analytic journals, 2013–2015. *Behavior Analysis: Research and Practice*, *17*, 84–91. doi.org/10.1037/bar0000071

Libby, M. E., Weiss, J. S., Bancroft, S., & Ahearn, W. H. (2008). A comparison of most-to-least and least-to-most prompting on the acquisition of solitary play skills. *Behavior Analysis in Practice*, *1*(1), 37–43.

Lieving, G. A., & Lattal, K. A. (2003). Recency, repeatability, and reinforcer retrenchment: An experimental analysis of resurgence. *Journal of the Experimental Analysis of Behavior*, *80*, 217–233. doi.org/10.1901/jeab.2003.80-217

Lilienfeld, S. O., Marshall, J., Todd, J. T., & Shane, H. C. (2014). The persistence of fad interventions in the face of negative scientific evidence: Facilitated communication for autism as a case example. *Evidence-based Communication Assessment and Intervention*, *8*(2), 62–101. doi:10.1080/17489539.2014.976332

Lindberg, J. S., Iwata, B. A., Kahng, S. W., & DeLeon, I. G. (1999). DRO contingencies: Analysis of variable-momentary schedules. *Journal of Applied Behavior Analysis*, *32*, 123–136.

Lindberg, J. S., Iwata, B. A., Roscoe, E. M., Worsdell, A. S., & Hanley, G. P. (2003). Treatment efficacy of noncontingent reinforcement during brief and extended application. *Journal of Applied Behavior Analysis*, *36*, 1–19.

Lindsley, O. R. (1956). Operant conditioning methods applied to research in chronic schizophrenia. *Psychiatric Research Reports*, *5*, 118–139.

Lindsley, O. R., & Skinner, B. F. (1954). A method for the experimental analysis of the behavior of psychotic patients. *American Psychologist*, *9*, 419–420,

Linehan, M. (1977). Issues in behavioral interviewing. In J. D. Cone & R. P. Hawkins (Eds.), *Behavioral assessment: New directions in clinical psychology* (pp. 30–51). New York: Bruner/Mazel.

Linscheid, T. R., & Meinhold, P. (1990). The controversy over aversives: Basic operant research and the side effects of punishment. In A. C. Repp & N. N. Singh (Eds.), *Perspectives on the use of nonaversive and aversive interventions for persons with developmental disabilities* (pp. 59–72). Sycamore, IL: Sycamore.

Linscheid, T. R., Iwata, B. A., Ricketts, R. W., Williams, D. E., & Griffin, J. C. (1990). Clinical evaluation of the self-injurious behavior inhibiting system (SIBIS). *Journal of Applied Behavior Analysis*, *23*, 53–78.

Lipinski, D. P., Black, J. L., Nelson, R. O., & Ciminero, A. R. (1975). Influence of motivational variables on the reactivity and reliability of self-recording. *Journal of Consulting and Clinical Psychology*, *43*, 637–646.

Lipkins, G., Hayes, S. C., & Hayes, L. J. (1993). Longitudinal study of derived stimulus relations in an infant. *Journal of Experimental Child Psychology*, *56*, 201–239.

Lipschultz, J., & Wilder, D. A. (2017). Recent research on the high-probability instructional sequence: A brief review. *Journal of Applied Behavior Analysis*, *50*, 424–428.

Litow, L., & Pumroy, D. K. (1975). A brief review of classroom group-oriented contingencies. *Journal of Applied Behavior Analysis*, *3*, 341–347.

Litt, M. D., & Schreibman, L. (1981). Stimulus-specific reinforcement in the acquisition of receptive labels by autistic children. *Analysis and Intervention in Developmental Disabilities*, *1*, 171–186.

Little, S. G., Akin-Little, A., & O'Neill, K. (2015). Group contingency interventions with children—1980–2010: A meta-analysis. *Behavior Modification*, *39*(2), 322–341.

Livingston, C. E., & Graff, R. B. (2018). Further evaluation of the use of preference categories to identify novel reinforcers: A systematic replication. *Behavioral Interventions*, *33*, 173–184. doi:10.1002/bin.1519

Lloyd, J. W., Eberhardt, M. J., & Drake, G. P., Jr. (1996). Group versus individual reinforcement contingencies with the context of group study conditions. *Journal of Applied Behavior Analysis*, *29*, 189–200.

Lo, Y. (2003). *Functional assessment and individualized intervention plans: Increasing the behavioral adjustment of urban learners in general and special education settings*. Doctoral dissertation, The Ohio State University, Columbus, OH.

Lo, Y., & Konrad, M. (2007). A field-tested task analysis for creating single-subject graphs using Microsoft Office Excel. *Journal of Behavioral Education*, *16*, 155–189.

Lo, Y. Y., Correa, V. I., & Anderson, A. L. (2015). Culturally responsive social skill instruction for Latino male students. *Journal of Positive Behavior Interventions*, *17*, 15–27.

Logan, K. R., & Gast, D. L. (2001). Conducing preference assessments and reinforcer testing for individuals with profound multiple disabilities: Issues and Procedures. *Exceptionality*, *9*(3), 123–134.

Logan, K. R., Jacobs, H. A., Gast, D. L., Smith, P. D., Daniel, J., & Rawls, J. (2001). Preferences and reinforcers for students with profound multiple disabilities: Can we identify them? *Journal of Developmental and Physical Disabilities*, *13*, 97–122.

Lombard, D., Neubauer, T. E., Canfield, D., & Winett, R. A. (1991). Behavioral community intervention to reduce the risk of skin cancer. *Journal of Applied Behavior Analysis*, *2*, 677–686.

Long, E. S., Miltenberger, R. G., Ellingson, S. A., & Ott, S. M. (1999). Augmenting simplified habit reversal in the treatment of oral-digital habits exhibited by individuals with mental retardation. *Journal of Applied Behavior Analysis*, *32*, 353–365.

Lotfizadeh, A. D., Edwards, T. L., Redner, R., & Poling, A. (2012). Motivating operations affect stimulus control: A largely overlooked phenomenon in discrimination learning. *The Behavior Analyst*, *35*, 89–100.

Lovaas, O. I. (1977). *The autistic child: Language development through behavior modification*. New York: Irvington.

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology*, *55*, 3–9.

Lovaas, O. I. (2003). *Teaching individuals with developmental delays*. Austin, TX: PRO-ED.

Lovaas, O. I., Koegel, R. L., & Schreibman, L. (1979). Stimulus overselectivity in autism: A review of research. *Psychological Bulletin*, *86*, 1236–1254. doi:10.1037/0033-2909.86.6.1236

Lovett, S., Rehfeldt, R. A., Garcia, Y., & Dunning, J. (2011). Comparison of a stimulus equivalence protocol and traditional lecture for teaching single-subject research designs. *Journal of Applied Behavior Analysis*, *44*, 819–833.

Lovitt, T. C. (1973). Self-management projects with children with behavioral disabilities. *Journal of Learning Disabilities*, *6*, 138–150.

Lovitt, T. C., & Curtiss, K. A. (1968). Effects of manipulating an antecedent event on mathematics response rate. *Journal of Applied Behavior Analysis*, *1*, 329–333.

Lovitt, T. C., & Curtiss, K. A. (1969). Academic response rates as a function of teacher and self-imposed contingencies. *Journal of Applied Behavior Analysis*, *2*, 49–53.

Lowe, C. F., Horne, P. J., Harris, F. D. A., & Randle, V. R. L. (2002). Naming and categorization in young children: Vocal tact training. *Journal of the Experimental Analysis of Behavior*, *78*, 527–249.

Lowe, C. F., Horne, P. J., & Hughes, J. C. (2005). Naming and categorization in young children: III. Vocal tact training and transfer of function. *Journal of the Experimental Analysis of Behavior*, *83*, 47–65. doi.org/10.1901/jeab.2005.31-04

Lowenkron, B. (1984). Coding responses and the generalization of matching to sample in children. *Journal of the Experimental Analysis of Behavior*, *42*, 1–18.

Lowenkron, B. (1988). Generalization of delayed identity matching in retarded children. *Journal of the Experimental Analysis of Behavior*, *50*, 163–172. doi:10.1901/jeab.1988.50-163

Lowenkron, B. (1989). Instructional control of generalized relational matching to sample in children. *Journal of the Experimental Analysis of Behavior*, *52*, 293–309. 10.1901/jeab.1989.52-293

Lowenkron, B. (1991). Joint control and the generalization of selection-based verbal behavior. *The Analysis of Verbal Behavior*, *9*, 121–126.

Lowenkron, B. (1998). Some logical functions of joint control. *Journal of the Experimental Analysis of Behavior*, *69*, 327–354.

Lowenkron, B. (2006). Joint control and the selection of stimuli from their description. *The Analysis of Verbal Behavior*, *22*, 129–150.

Luce, S. C., & Hall, R. V. (1981). Contingent exercise: A procedure used with differential reinforcement to reduce bizarre verbal behavior. *Education and Treatment of Children*, *4*, 309–327.

Luce, S. C., Delquadri, J., & Hall, R. V. (1980). Contingent exercise: A mild but powerful procedure for suppressing inappropriate verbal and aggressive behavior. *Journal of Applied Behavior Analysis*, *13*, 583–594.

Luciano, C., Becerra, I. G., & Valverde, M. R. (2007). The role of multiple-exemplar training and naming in establishing derived equivalence in an infant. *Journal of the Experimental Analysis of Behavior*, *87*, 349–365.

Luciano, C., Rodriguez, M., Manas, I., Ruiz, F., Berens, N. M., & Valdivio-Salas, S. (2009). Acquiring the earliest relational operants: Coordination, distinction, opposition, comparison, and hierarchy. In R. A. Rehfeldt & Y. Barnes-Holmes (Eds.), *Derived relational responding: Applications for learners with autism and other developmental disabilities: A progressive guide to change* (pp. 149–172). New Oakland, CA: Harbinger Publications.

Luczynski, K. C., & Hanley, G. P. (2009). Do children prefer contingencies? An evaluation of the efficacy of and preference for contingent versus noncontingent social reinforcement during play. *Journal of Applied Behavior Analysis*, *42*, 511–525. doi:10.1901/jaba.2009.42-511

Ludwig, R. L. (2004). *Smiley faces and spinners: Effects of self-monitoring of productivity with an indiscriminable contingency of reinforcement on the on-task behavior and academic productivity by kindergarteners during independent seatwork*. Unpublished master's thesis, The Ohio State University, Columbus, OH.

Luffman, W. (2012). *Using class-specific reinforcers and responses with conditional discrimination training and recombinative generalization with children*. Unpublished master's thesis, University of North Carolina Wilmington.

Luinge, M. R., Post, W. J., Wit, H. P., & Goorhuis-Brouwer, S. M. (2006). The ordering of milestones in language development for children from 1 to 6 years of age. *Journal of Speech, Language, and Hearing Research*, *49*, 923–940.

Luiselli, J. K. (1984). Controlling disruptive behaviors of an autistic child: Parent-mediated contingency management in the home setting. *Education and Treatment of Children*, *3*, 195–203.

Luiselli, J. K., Woods, K. E., & Reed, D. D. (2011). Review of sports performance research with youth, collegiate, and elite athletes. *Journal of Applied Behavior Analysis*, *44*, 999–1002.

Luke, M., & Alavosius, M. (2011). Adherence with universal precautions after immediate, personalized performance feedback. *Journal of Applied Behavior Analysis*, *44*, 967–971.

Luke, N. R., Greer, R. D., Singer-Dudek, J., & Keohane, D. (2011). The emergence of autoclitic frames in atypically and typically developing children as a function of multiple exemplar instruction. *The Analysis of Verbal Behavior*, *27*, 141–156.

Lumley, V., Miltenberger, R., Long, E., Rapp, J., & Roberts, J. (1998). Evaluation of a sexual abuse prevention program for adults with mental retardation. *Journal of Applied Behavior Analysis*, *31*, 91–101.

Lundgren, T., Dahl, J., & Hayes, S. C. (2008). Evaluation of mediators of change in the treatment of epilepsy with acceptance and commitment therapy. *Journal of Behavioral Medicine*, *31*, 225–235.

Luyben, P. D. (2009). Applied behavior analysis: Understanding and changing behavior in the community—A representative review. *Journal of Prevention & Intervention Community 37*(3), 230–253. doi:10.1080/10852350902975884

Ly, K. H., Asplund, K., & Andersson, G. (2014). Stress management for middle managers via an acceptance and commitment-based smartphone application: A randomized controlled trial. *Internet Interventions*, *1*, 95–101.

Lydon, S., Healy, O., O'Reilly, M. F., & Lang, R. (2012). Variations in functional analysis methodology: A systematic review. *Journal of Developmental and Physical Disabilities*, *24*, 301–326.

Lynch, D., & Keenan, M. (2018). The good behavior game: Maintenance effects. *Journal of International Educational Research*, *87*, 91–99. doi.org/10.1016/j.ijer.2016.05.005

Lynch, D. C., & Cuvo, A. J. (1995). Stimulus equivalence instruction of fraction-decimal relations. *Journal of Applied Behavior Analysis*, *28*, 115–126.

Lynch, D. C., & Green, G. (1991). Development and crossmodal transfer on contextual control of emergent stimulus relations. *Journal of the Experimental Analysis of Behavior*, *56*, 139–154.

Lyon, C. S., & Lagarde, R. (1997). Tokens for success: Using the graduated reinforcement system. *Teaching Exceptional Children*, *29*(6), 52–57.

Ma, M. L., Miguel, C. F., & Jennings, A. M. (2016). Training intraverbal naming to establish equivalence class performances. *Journal of the Experimental Analysis of Behavior*, *105*, 409–426. doi:10.1002/jeab.203

Maag, J. W., Reid, R., & DiGangi, S. A. (1993). Differential effects of self-monitoring attention, accuracy, and productivity. *Journal of Applied Behavior Analysis*, *26*, 329–344.

MacCorquodale, K., & Meehl, P. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Record*, *55*, 95–107.

MacDonald, J., & Ahearn, W. H. (2015). Teaching observational learning to children with autism. *Journal of Applied Behavior Analysis*, *48*, 800–816.

MacDonald, J. M., Ahearn, W. H., Parry-Cruwys, D., & Bancroft, S. (2013). Persistence during extinction: Examining the effects of continuous and intermittent reinforcement on problem behavior. *Journal of Applied Behavior*, *46*, 333–338. doi:10.1002/jaba.3

MacDuff, G. S., Krantz, P. J., & McClannahan, L. E. (1993). Teaching children with autism to use photographic activity schedules: Maintenance and generalization of complex response chains. *Journal of Applied Behavior Analysis*, *26*, 89–97.

Mace, F. C. (1996). In pursuit of general behavioral relations. *Journal of Applied Behavior Analysis*, *29*, 557–563.

Mace, F. C., & Belfiore, P. (1990). Behavioral momentum in the treatment of escape-motivated stereotypy. *Journal of Applied Behavior Analysis*, *23*, 507–514.

Mace, F. C., & Critchfield, T. S. (2010). Translational research in behavior analysis: Historical traditions and imperative for the future. *Journal of the Experimental Analysis of Behavior*, *93*, 293–312.

Mace, F. C., Hock, M. L., Lalli, J. S., West, B. J., Belfiore, P., Pinter, E., & Brown, D. K. (1988). Behavioral momentum in the treatment of noncompliance. *Journal of Applied Behavior Analysis*, *21*, 123–141.

Mace, F. C., Lalli, J. S., Shea, M. C., & Nevin, J. A. (1988). Behavioral momentum in college basketball. *Journal of Applied Behavior Analysis*, *25*, 657–663.

Mace, F. C., McComas, J. J., Mauro, B. C., Progar, P. R., Taylor, B., Ervin, R., & Zangrillo, A. N. (2010). Differential reinforcement of alternative behavior increases resistance to extinction: Clinical demonstration, animal modeling, and clinical test of one solution. *Journal of the Experimental Analysis of Behavior*, *93*, 349–367. doi.org/10.1901/jeab.2010.93-349

Mace, F. C., Page, T. J., Ivancic, M. T., & O'Brien, S. (1986). Effectiveness of brief time-out with and without contingency delay: A comparative analysis. *Journal of Applied Behavior Analysis*, *19*, 79–86.

Mace, F. C., Pratt, J. L., Prager, K. L., & Pritchard, D. (2011). An evaluation of three methods of saying "No" to avoid an escalating response class hierarchy. *Journal of Applied Behavior Analysis*, *44*, 83–94.

MacNeil, J., & Thomas, M. R. (1976). Treatment of obsessive-compulsive hairpulling (trichotillomania) by behavioral and cognitive contingency manipulation. *Journal of Behavior Therapy and Experimental Psychiatry, 7*, 391–392.

MacSuga-Gage, A. S., & Simonsen, B. (2015). Examining the impact of opportunities to respond on student outcomes: A systematic review of the literature. *Education and Treatment of Children, 38*(2), 211–240.

Madden, G. J. (Ed.). (2013). *APA handbook of behavior analysis*. Washington, DC: American Psychological Association.

Madden, G. J., & Johnson, P. S. (2010). A discounting primer. In G. J. Madden & W. K. Bickel (Eds.), *Impulsivity: The behavioral and neurological science of discounting* (pp. 11–37). Washington, DC: American Psychological Association.

Madden, G. J., Chase, P. N., & Joyce, J. H. (1998). Making sense of sensitivity in the human operant literature. *The Behavior Analyst, 21*, 1–12.

Madden, G. J., Petry, N. M., Badger, G. J., & Bickel, W. K. (1997). Impulsive and self-control choices in opioid-dependent patients and non-drug-using control participants: Drug and monetary rewards. *Experimental and Clinical Psychopharmacology, 5*, 256–262.

Madsen, C. H., Becker, W. C., Thomas, D. R., Koser, L., & Plager, E. (1968). An analysis of the reinforcing function of "sit down" commands. In R. K. Parker (Ed.), *Readings in educational psychology* (pp. 265–278). Boston: Allyn & Bacon.

Maggin, D., Johnson, A., Chafouleas, S., Ruberto, L., & Berggren, M. (2012). A systematic evidence review of school-based group contingency interventions for students with challenging behavior. *Journal of School Psychology, 50*, 625–654.

Maggin, D. M., Briesh, A. M., Chafouleas, S. M., Ferguson, T. D., & Clark, C. (2014). A comparison of rubrics for identifying empirically supported practices with single-case research. *Journal of Behavioral Education, 25*, 287–311. doi 10.1007/s10864-013-9187-z

Maggin, D. M., O'Keefe, B. V., & Johnson, A. H. (2011). A quantitative synthesis of single subject meta-analysis in special education, 1985–2009. *Exceptionality, 19*, 109–135. doi:10.1080/09362835.2011.565725

Maglieri, K. A., DeLeon, I. G., Rodriguez-Catter, V., & Sevin, B. M. (2000). Treatment of covert food stealing in an individual with Prader-Willi syndrome. *Journal of Applied Behavior Analysis, 33*, 615–618.

Magoon, M.A., & Critchfield, T. S. (2008). Concurrent schedules of positive and negative reinforcement: Differential-impact and differential-outcomes effects. *Journal of the Experimental Analysis of Behavior, 90*, 1–22.

Magoon, M. A., Critchfield, T. S., Merrill, D., Newland, M. C., & Schneider, W. J. (2017). Are positive and negative reinforcement "different"? Insights from a free-operant differential outcomes effect. *Journal of the Experimental Analysis of Behavior, 107*, 39–64.

Maheady, L., Michielli-Pendl, J., Harper, G., & Mallette, B. (2006). The effects of numbered heads together with and without an incentive package on the science test performance of a diverse group of sixth graders. *Journal of Behavioral Education, 15*(1), 24–39. http://dx.doi.org/10.1007/s10864-005-9002-6

Mahoney, M. J. (1971). The self-management of covert behavior: A case study. *Behavior Therapy, 2*, 575–578.

Mahoney, M. J. (1976). Terminal terminology. *Journal of Applied Behavior Analysis, 9*, 515–517.

Maine Administrators of Services for Children with Disabilities (2000). *Report of the MADSEC Autism Task Force, Revised Edition*. Kennebec Centre, RR 2 Box 1856, Manchester, ME 04351, http://www.madsec.org/docs/atf.htm

Malanga, P. R. (2017). *Strategies for programming generalized responding*. Unpublished manuscript, p. 10. Used with permission.

Maletzky, B. M. (1974). Behavior recording as a treatment: A brief note. *Behavior Therapy, 5*, 107–111.

Maloney, K. B., & Hopkins, B. L. (1973). The modification of sentence structure and its relationship to subjective judgments of creativity in writing. *Journal of Applied Behavior Analysis, 6*, 425–433.

Malott, M. E. (2003). *Paradox of organizational change: Engineering organizations with behavioral systems analysis*. Oakland, CA: Context Press.

Malott, M. E. (2016). What studying leadership can teach us about the science of behavior. *The Behavior Analyst, 39*, 47–74. doi:10.1007/s40614-015-0049-y

Malott, R. W. (1981). *Notes from a radical behaviorist*. Kalamazoo, MI: Author.

Malott, R. W. (1984). Rule-governed behavior, self-management, and the developmentally disabled: A theoretical analysis. *Analysis and Intervention in Developmental Disabilities, 6*, 53–68.

Malott, R. W. (1988). Rule-governed behavior and behavioral anthropology. *The Behavior Analyst, 11*, 181–203.

Malott, R. W. (1989). The achievement of evasive goals: Control by rules describing contingencies that are not direct-acting. In S. C. Hayes (Ed.), *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 269–322). New York: Plenum.

Malott, R. W. (2003). Behavior analysis and linguistic productivity. *The Analysis of Verbal Behavior, 19*, 11–18.

Malott, R. W. (2005a). Self-management. In M. Hersen & J. Rosqvist (Eds.), *Encyclopedia of behavior modification and cognitive behavior therapy (Vol. I: Adult clinical applications)* (pp. 516–521). Newbury Park, CA: Sage.

Malott, R. W. (2005b). Behavioral systems analysis and higher education. In W. L. Heward, T. E. Heron, N. A. Neef, S. M. Peterson, D. M. Sainato, G. Cartledge, R. Gardner III, L. D. Peterson, S. B. Hersh, & J. C. Dardig (Eds.), *Focus on behavior analysis in education: Achievements, challenges, and opportunities* (pp. 211–236). Upper Saddle River, NJ: Merrill/Prentice Hall.

Malott, R. W. (2012). *I'll stop procrastinating when I get around to it: Plus other cool ways to succeed in school and life using behavior analysis to get your act together*. Kalamazoo, MI: Behaviordelia.

Malott, R. W., & Shane, J. T. (2014). *Principles of behavior* (7th ed.). New York: Routledge.

Malott, R. W., General, D. A., & Snapper, V. B. (1973). *Issues in the analysis of behavior*. Kalamazoo, MI: Behaviordelia.

Malott, R. W., Hartlep, P., Hartlep, S., Vaughan, P., & Smith, J. E. (1971). *Contingency management in education and other equally exciting places*. Kalamazoo, MI: Behaviordelia.

Malott, R. W., Tillema, M., & Glenn, S. (1978). *Behavior analysis and behavior modification: An introduction*. Kalamazoo, MI: Behaviordelia.

Mancuso, C., & Miltenberger, R. G. (2016). Using habit reversal to decrease filled pauses in public speaking. *Journal of Applied Behavior Analysis, 49*, 188–192. doi:10.1002/jaba.267

Mank, D. M., & Horner, R. H. (1987). Self-recruited feedback: A cost-effective procedure for maintaining behavior. *Research in Developmental Disabilities, 8*, 91–112.

Mann, M., & Toles, T. (2016). *The madhouse effect: How climate change denial is threatening our planet, destroying our politics, and driving us crazy*. New York: Columbia University Press.

Manolov, R., & Solanas, A. (2013). A comparison of mean phase difference and generalized least squares for analyzing single-case data. *Journal of School Psychology, 51*(2), 201–215. doi:10.1016/j.jsp.2012.12.005

Maple, T. L., & Segura, V. D. (2015). Advancing behavior analysis in zoos and aquariums. *The Behavior Analyst, 38*, 77–91.

March, R. E., Horner, R. H., Lewis-Palmer, T., Brown, D., Crone, D, Todd, A. W., & Carr, E. (2000). *The Functional Assessment Checklist for Teachers (FACT)*. Educational and Community Supports, Eugene, OR.

Marcott, S. A., Shakun, J. D., Clark, P. U., & Mix, A. C. (2013). A reconstruction of regional and global temperature for the past 11,300 years. *Science, 339*(6124), 1198–1201.

Marcus, B. A., & Vollmer, T. R. (1995). Effects of differential negative reinforcement on disruption and compliance. *Journal of Applied Behavior Analysis, 28*, 229–230.

Marholin, D., II, & Steinman, W. (1977). Stimulus control in the classroom as a function of the behavior reinforced. *Journal of Applied Behavior Analysis, 10*, 465–478.

Marholin, D., Touchette, P. E., & Stuart, R. M. (1979). Withdrawal of chronic chlorpromazine medication: An experimental analysis. *Journal of Applied Behavior Analysis, 12*, 150–171.

Markle, S. M. (1962). *Good frames and bad: A grammar of frame writing* (2nd ed.). New York: Wiley.

Markwardt, F. C. (1989). *Peabody Individual Achievement Test*. Circle Pines, MN: American Guidance Service.

Marlatt, G. A., & Kristeller, J. L. (1999). Mindfulness and meditation. In W. R. Miller (Ed.), *Integrating spirituality into treatment* (pp. 67–84). Washington, DC: American Psychological Association.

Marmolejo, E. K., Wilder, D. A., & Bradley, L. (2004). A preliminary analysis of the effects of response cards on student performance and participation in

an upper division university course. *Journal of Applied Behavior Analysis*, *37*, 405–410.

Marr, J. (2003). Empiricism. In K. A. Lattal & P. C. Chase (Eds.), *Behavior theory and philosophy* (pp. 63–82). New York: Kluwer/Plenum.

Marr, M. J. (2003). The stitching and the unstitching: What can behavior analysis have to say about creativity? *The Behavior Analyst*, *26*(1), 15–27.

Marr, M. J. (2006). Through the looking glass: Symmetry in behavioral principles? *The Behavior Analyst*, *29*, 125–128.

Marr, M. J. (2017). The future of behavior analysis: Foxes and hedgehogs revisited. *The Behavior Analyst*, *40*, 197–207. http://dx.doi.org/10.1007/s40614-017-0107-8

Marshall, A. E., & Heward, W. L. (1979). Teaching self-management to incarcerated youth. *Behavioral Disorders*, *4*, 215–226.

Marshall, K. J., Lloyd, J. W., & Hallahan, D. P. (1993). Effects of training to increase self-monitoring accuracy. *Journal of Behavioral Education*, *3*, 445–459.

Martella, R., Leonard, I. J., Marchand-Martella, N. E., & Agran, M. (1993). Self-monitoring negative statements. *Journal of Behavioral Education*, *3*, 77–86.

Martens, B. K., DiGennaro, F. D., Reed, D. D., Szczech, F. M., & Rosenthal, B. D. (2008). Contingency space analysis: An alternative method for identifying contingent relations from observational data. *Journal of Applied Behavior Analysis*, *41*, 69–81.

Martens, B. K., Gertz, L. E., Werder, C. S., Rymanowski, J. L., & Shankar, K. H. (2014). Measures of association in contingency space analysis. *Journal of Mathematical Psychology*, *59*, 114–119.

Martens, B. K., Lochner, D. G., & Kelly, S. Q. (1992). The effects of variable-interval reinforcement on academic engagement: A demonstration of matching theory. *Journal of Applied Behavior Analysis*, *25*, 143–151.

Martens, B. K., Witt, J. C., Elliott, S. N., & Darveaux, D. (1985). Teacher judgments concerning the acceptability of school-based interventions. *Professional Psychology: Research and Practice*, *16*, 191–198.

Martin, G., & Pear, J. (2003). *Behavior modification: What it is and how to do it* (7th ed.). Englewood Cliffs, NJ: Prentice Hall.

Martin, G., & Pear, J. (2015). *Behavior modification: What it is and how to do it* (10th ed.). New York: Routledge.

Martinez-Diaz, J. A. (2003). *Raising the bar*. Presidential address presented at the annual conference of the Florida Association for Behavior Analysis, St. Petersburg, FL.

Martins, L. L., Hübner, M. C., Gomes, F. P., Pinto Portugal, M., & Treu, K. E. (2015). Effect of the qualifying autoclitic 'is' in conditional discrimination training and equivalence tests. *Acta Colombiana De Psicología*, *18*, 37–46.

Marzullo-Kerth, D., Reeve, S. A., Reeve, K. F., & Townsend, D. B. (2011). Using multiple-exemplar training to teach a generalized repertoire of sharing to children with autism. *Journal of Applied Behavior Analysis*, *38*, 279–294.

Mastellone, M. (1974). Aversion therapy: A new use of the old rubberband. *Journal of Behavior Therapy and Experimental Psychiatry*, *5*, 311–312.

Matos, M. A., & Hubner-d'Oliveira, M. M. (1992). Equivalence relations and reading. In Hayes, S. C., & Hayes, L. J. (Eds.), *Understanding verbal relations* (pp. 83–94). Reno, NV: Context Press.

Matos, M. A., Avanzi, A. L., & McIlvane, W. J. (2006). Rudimentary reading repertoires via stimulus equivalence and recombination of minimal units. *The Analysis of Verbal Behavior*, *22*, 3–19.

Matson, J. L., & Taras, M. E., (1989). A 20-year review of punishment and alternative methods to treat problem behaviors in developmentally delayed persons. *Research in Developmental Disabilities*, *10*, 85–104.

Mattaini, M. A. (1995). Contingency diagrams as teaching tools. *The Behavior Analyst*, *18*, 93–98.

Matthews, B. A., Shimoff, E., & Catania, A. C. (1987). Saying and doing: A contingency-space analysis. *Journal of Applied Behavior Analysis*, *20*, 69–74.

Matute, H., & Miller, R. R. (1998). Detecting causal relations. In W. H. O'Donohue (Ed.). *Learning and behavior therapy* (pp. 483–497). Boston: Allyn and Bacon.

Matyas, T. A., & Greenwood, K. M. (1990). Visual analysis of single-case time series: Effects of variability, serial dependence, and magnitude of intervention effects. *Journal of Applied Behavior Analysis*, *23*, 341–351.

Maurice, C. (1993). *Let me hear your voice: A family's triumph over autism*. New York: Fawcett Columbine.

Maurice, C. (2017). The autism wars. In W. L. Heward, S. R. Alber-Morgan, & M. Konrad, *Exceptional children: An introduction to special education* (11th ed.) (pp. 266–267). Upper Saddle River, NJ: Pearson.

Maurice, C., Green, G., & Luce, S.C. (1996). *Behavior interventions for young children with autism*. Austin, TX: PRO-ED.

Maxwell, J. C. (2003). *There's no such thing as "business" ethics: There's only one rule for decision-making*. New York: Warner Business Books: A Time Warner Company.

May, R. J., Downs, R., Marchant, A., & Dymond, S. (2016). Emergent verbal behavior in preschool children learning a second language. *Journal of Applied Behavior Analysis*, *49*, 711–716. doi:10.1002/jaba.301

May, R. J., Stewart, I., Baez, L., Freegard, G., & Dymond, S. (2017). Arbitrarily applicable spatial relational responding. *Journal of the Experimental Analysis of Behavior*, *107*, 234–257.

Maydak, M., Stromer, R., Mackay, H. A., & Stoddard, L. T. (1995). Stimulus classes in matching to sample and sequence production: The emergence of numeric relations. *Research in Developmental Disabilities*, *16*, 179–204.

Mayer, G. R., Butterworth, T., Nafpaktitis, M., & Sulzer-Azaroff, B. (1983). Preventing school vandalism and improving discipline: A three-year study. *Journal of Applied Behavior Analysis*, *16*, 355–369.

Mayer, G. R., Sulzer, B., & Cody, J. J. (1968). The use of punishment in modifying student behavior. *Journal of Special Education*, *2*, 323–328.

Mays, N. M., & Heflin, L. J. (2011). Increasing independence in self-care tasks for children with autism using self-operated auditory prompts. *Research in Autism Spectrum Disorders*, *5*(4), 1351–1357.

Mazaleski, J. L., Iwata, B. A., Rogers, T. A., Vollmer, T. R., & Zarcone, J. R. (1994). Protective equipment as treatment for stereotypic hand mouthing: Sensory extinction or punishment effects? *Journal of Applied Behavior Analysis*, *27*, 345–355.

Mazur, J. E., & Fantino, E. (2014). Choice. In F. K. McSweeney & E. S. Murphy (Eds.), *The Wiley-Blackwell handbook of operant and classical conditioning* (pp. 195–220). Oxford, UK: John Wiley & Sons.

Mazur, J. E., & Logue, A. W. (1978). Choice in a "self-control" paradigm: Effects of a fading procedure. *Journal of the Experimental Analysis of Behavior*, *30*, 11–17.

McAdam, D. B., Klatt, K. P., Koffarnus, M., Dicesare, A., Solberg, K., Welch, C., & Murphy, S. (2005). The effects of establishing operations on preferences for tangible items. *Journal of Applied Behavior Analysis*, *38*, 107–110.

McAdam, D. B., O'Cleirigh, C. M., & Cuvo, A. J. (1993). Self-monitoring and verbal feedback to reduce stereotypic body rocking in a congenitally blind adult. *RE:view: Rehabilitation and Education for Blindness and Visual Impairment*, *24*, 163–172.

McAllister, L. W., Stachowiak, J. G., Baer, D. M., & Conderman, L. (1969). The application of operant conditioning techniques in a secondary school classroom. *Journal of Applied Behavior Analysis*, *2*, 277–285.

McClannahan, L. E., & Krantz, P. J. (2010). *Activity schedules for children with autism: Teaching independent behavior* (2nd ed.). Bethesda, MD: Woodbine House.

McClannahan, L. E., MacDuff, G. S., & Krantz, P. J. (2002). Behavior analysis and intervention for adults with autism. *Behavior Modification*, *26*, 9–27.

McCollow, M. M., Curiel, E. S. L., Davis, C. A., & Sainato, D. (2016). Utilizing antecedent strategies in early childhood settings. *Young Exceptional Children*, *14*(4), 5–19.

McComas, J., Hoch, H., Poane, D., & El-Roy, D. (2000). Escape behavior during academic tasks: A preliminary analysis of idiosyncratic establishing operations. *Journal of Applied Behavior Analysis*, *33*, 479–493.

McConachie, D. A. J., McKenzie, K., Morris, P. G., & Walley, R. M. (2014). Acceptance and mindfulness-based stress management for support staff caring for individuals with intellectual disabilities. *Research in Developmental Disabilities*, *35*, 1216–1227.

McConnell, M. E. (1999). Self-monitoring, cueing, recording, and managing: Teaching students to manage their own behavior. *Teaching Exceptional Children*, *32*(2), 14–21.

McCord, B. E., Iwata, B. A., Galensky, T. L., Ellingson, S. A., & Thomson, R. J. (2001). Functional analysis and treatment of problem behavior evoked by noise. *Journal of Applied Behavior Analysis*, *34*, 447–462.

McCurdy, B. L., Lannie, A. L., & Barnabas, E. (2009). Reducing disruptive behavior in an urban school cafeteria: An extension of the Good Behavior Game. *Journal of School Psychology*, *47*, 39–54.

McDonagh, E. C., McIlvane, W. J., & Stoddard, L. T. (1984). Teaching coin equivalences via matching to sample. *Applied Research in Mental Retardation*, *5*, 177–197.

McDonough, M. R., Johnson, A. K., & Waters, B. J. (2017). Reducing attention maintained bizarre speech in an adult woman. *Behavior Analysis: Research and Practice*, *17*(2), 197–203. doi.org/10.1037/bar0000059

McDougall, D. (2005). The range-bound changing criterion design. *Behavioral Interventions*, *20*, 129–137.

McDougall, D., Heine, R. C., Wiley, L. A., Sheehey, M. D., Skanashi, K. K., Cook, B. G., & Cook, L. (2017). Meta-analysis of behavioral self-management techniques used by students with disabilities in inclusive settings. *Behavioral Interventions*, 1–19.

McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation*, *97*, 359–372.

McEntee, J. E., & Saunders, R. R. (1997). A response-restriction analysis of stereotypy in adolescents with mental retardation: Implications for applied behavior analysis. *Journal of Applied Behavior Analysis*, *30*, 485–506.

McEnteggart, C. (2018). A brief tutorial on Acceptance and Commitment Therapy as seen through the lens of derived stimulus relations. *Perspectives on Behavior Science*, *41*, 215–227.

McEvoy, M. A., & Brady, M. P. (1988). Contingent access to play materials as an academic motivator for autistic and behavior disordered children. *Education and Treatment of Children*, *11*, 5–18.

McFall, R. M. (1977). Parameters of self-monitoring. In R. B. Stuart (Ed.), *Behavioral self-management* (pp. 196–214). New York: Bruner/Mazel.

Mcgee, G. G., Krantz, P. J., & McClannahan, L. E. (1985). The facilitative effects of incidental teaching on preposition use by autistic children. *Journal of Applied Behavior Analysis*, *18*, 17–31.

Mcgee, G. G., Morrier, M., & Daly, T. (1999). An incidental teaching approach to early intervention for toddlers with autism. *Journal of the Association for Persons with Severe Handicaps*, *24*, 133–146.

McGhan, A. C., & Lerman, D. C. (2013). An assessment of error-correction procedures for learners with autism. *Journal of Applied Behavior Analysis*, *46*, 626–639.

McGill, P. (1999). Establishing operations: Implications for the assessment, treatment, and prevention of problem behavior. *Journal of Applied Behavior Analysis*, *32*, 393–418.

McGoey, K. E., & DuPaul, G. J. (2000). Token reinforcement and response cost procedures: Reducing the disruptive behavior of preschool children with attention-deficit/hyperactivity disorder. *School Psychology Quarterly*, *15*(3), 330–343. http://dx.doi.org/10.1037/h0088790

McGreevy, P., Fry, T., & Cornwall, C. (2012). *Essential for living: A communication, behavior, and functional skills assessment, curriculum, and teaching manual*. Winter Park, FL: Patrick McGreevy Publishing.

McGuffin, M. E., Martz, S. A., & Heron, T. E. (1997). The effects of self-correction versus traditional spelling on the spelling performance and maintenance of third grade students. *Journal of Behavioral Education*, *7*, 463–476.

McGuire, M. T., Wing, R. R., Klem, M. L., & Hill, J. O. (1999). Behavioral strategies of individuals who have maintained long-term weight losses. *Obesity Research*, *7*, 334–341.

McHugh, L., Barnes-Holmes, Y., & Barnes-Holmes, D. (2004). Perspective-taking as relational responding: A developmental profile. *The Psychological Record*, *54*, 115–144.

McHugh, L., Barnes-Holmes, Y., & Barnes-Holmes, D. (2009). Understanding and training perspective taking as relational responding. In R. A. Rehfeldt & Y. Barnes-Holmes (Eds.). *Derived relational responding: Applications for learners with autism and other developmental disabilities: A progressive guide to change* (pp. 281–300). New Oakland, CA: Harbinger Publications.

McHugh, L., Barnes-Holmes, Y., Barnes-Holmes, D., Stewart, I., & Dymond, S. (2007). Deictic relational complexity and the development of deception. *The Psychological Record*, *57*, 517–531.

McIlvane, W. J. (2009). Translational behavior analysis: From laboratory science in stimulus control to intervention with persons with neurodevelopmental disabilities. *The Behavior Analyst*, *32*, 273–280.

McIlvane, W. J., & Dube, W. V. (1992). Stimulus control shaping and stimulus control topographies. *The Behavior Analyst*, *15*, 89–94.

McIlvane, W. J., & Dube, W. V. (2003). Stimulus control topography coherence theory: Foundations and extensions. *The Behavior Analyst*, *26*, 195–213.

McIlvane, W. J., & Stoddard, L. T. (1981). Acquisition of matching-to-sample performances in severe retardation: Learning by exclusion. *Journal of Mental Deficiency Research*, *25*, 33–48.

McIlvane, W. J., Dube, W. V., Green, G., & Serna, R. W. (1993). Programming conceptual and communication skill development: A methodological stimulus class analysis. In A. P. Kaiser & D. B. Gray (Eds.), *Enhancing children's communication* (Vol. 2, pp. 243–285). Baltimore: Brookes.

McIntyre, L., Gresham, F. M., DiGennaro, F. D., & Reed, D. D. (2007). Treatment integrity of school-based interventions with children in the *Journal of Applied Behavior Analysis* 1991–2005. *Journal of Applied Behavior Analysis*, *40*, 659–672.

McKenzie, S. D., Smith, R. G., Simmons, J. N., & Soderlund, M. J. (2008). Using a stimulus correlated with reprimands to suppress automatically maintained eye poking. *Journal of Applied Behavior Analysis*, *41*, 255–259.

McKenzie, T. L. (2016). Context matters: Systematic observation of place-based physical activity. *Research Quarterly for Exercise and Sport*, *87*, 334–341. doi:10.1080/02701367.2016.1234302

McKenzie, T. L., & Cohen, D. A. (2006). *SOPARC (System for Observing Play and Recreation in Communities): Description and Procedure Manual*. Retrieved December 17, 2017, from https://activelivingresearch.org/sites/default/files/SOPARC_Protocols.pdf

McKenzie, T. L., Marshall, S. J., Sallis, J. F., & Conway, T. L. (2000). Leisure-time physical activity in school environments: An observational study using SOPLAY. *Preventive Medicine*, *30*, 70–77.

McKenzie, T. L., Sallis, J. F., & Rosengard, P. (2009). Beyond the stucco tower: Design, development, and dissemination of the SPARK physical education programs. *Quest*, *61*, 114–127.

McKerchar, P. M., & Thompson, R. H. (2004). A descriptive analysis of potential reinforcement contingencies in the preschool classroom. *Journal of Applied Behavior Analysis*, *37*, 431–444.

McLaughlin, D. M., & Carr, E. G. (2005). Quality of rapport as a setting event for problem behavior: Assessment and intervention. *Journal of Positive Behavior Interventions*, *7*, 68–91.

McLaughlin, T., & Malaby, J. (1972). Reducing and measuring inappropriate verbalizations in a token classroom. *Journal of Applied Behavior Analysis*, *5*, 329–333.

McNeish, J., Heron, T., & Okyere, B. (1992). Effects of self-correction on the spelling performance of junior high students with learning disabilities. *Journal of Behavioral Education*, *2*(1), 17–27.

McSweeney, F. K., & Murphy, E. S. (Eds.). (2014). *The Wiley-Blackwell handbook of operant and classical conditioning*. Oxford, UK: John Wiley & Sons.

McWilliams, R., Nietupski, J., & Hamre-Nietupski, S. (1990). Teaching complex activities to students with moderate handicaps through the forward chaining of shorter total cycle response sequences. *Education and Training in Mental Retardation*, *25*(3), 292–298.

Mechling, L. (2007). Assistive technology as a self-management tool for prompting students with intellectual disabilities to initiate and complete daily tasks: A literature review. *Education and Training in Developmental Disabilities*, *42*, 253–269.

Mechling, L. C. (2006). Comparison of the effects of three approaches on the frequency of stimulus activations via a single switch by students with profound intellectual disabilities. *Journal of Special Education*, *40*, 94–102.

Mechling, L. C. (2011). Review of twenty-first century portable electronic devices for persons with moderate intellectual disabilities and autism spectrum disorders. *Education and Training in Autism and Developmental Disabilities*, *46*, 476–498.

Mechling, L., Ayres, K. M., Purrazzella, K., & Purrazzella, K. (2012). Evaluation of the performance of fine and gross motor skills within multi-step tasks by adults with moderate intellectual disability when using video models. *Journal of Developmental and Physical Disabilities*, *24*, 469–486. doi 10.1007/s10882-012-9284-2

Mechner, F. (1959). A notation system for the description of behavioral procedures. *Journal of the Experimental Analysis of Behavior*, *2*, 133–150.

Mechner, F. (2017). A behavioral and biological analysis of aesthetics: Implications for research and applications. *The Psychological Record*, *68*. doi.org/10.1007/s40732-017-0228-1

Meichenbaum, D., & Goodman, J. (1971). The developmental control of operant motor responding by verbal operants. *Journal of Experimental Child Psychology, 7*, 553–565.

Meindl, J. N., Ivy, J. W., Miller, N., Neef, N. A., & Williamson, R. L. (2013). An examination of stimulus control in fluency-based strategies: SAFMEDS and generalization. *Journal of Behavioral Education, 22*, 229–252.

Melchiori, L .E., de Souza, D. G., & de Rose, J. C. (2000). Reading, equivalence, and recombination of units: A replication with students with different learning histories. *Journal of Applied Behavior Analysis, 33*, 97–100.

Mellalieu, S. D., Hanton, S., & O'Brien, M. (2006). The effects of goal setting on rugby performance. *Journal of Applied Behavior Analysis, 39*, 257–261.

Melzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198*, 75–78.

Merbitz, C., Morris, J., & Grip, J. C. (1989). Ordinal scales and foundations of misinference. *Archives of Physical Medicine and Rehabilitation, 70*, 308–312.

Merbitz, C. T., Merbitz, N. H., & Pennypacker, H. S. (2015). *The Behavior Analyst*, 39, 333–338. doi:10.1007/s40614-015-0048-z

Mercatoris, M., & Craighead, W. E. (1974). Effects of nonparticipant observation on teacher and pupil classroom behavior. *Journal of Educational Psychology, 66*, 512–519.

Meredith, S. E., Grabinski, M. J., & Dallery, J. (2011). Internet-based group contingency management to promote abstinence from cigarette smoking: A feasibility study. *Drug and Alcohol Dependence, 118*, 23–30.

Metz, E. R., Kohn, C. S., Schultz, N. R., & Bettencourt, K. (2017). Evaluation of pour training procedures for college students. *Behavior Analysis: Research and Practice, 17*(1), 18–32. doi.org/10.1037/bar0000038

Michael, J. (1974). Statistical inference for individual organism research: Mixed blessing or curse? *Journal of Applied Behavior Analysis, 7*, 647–653.

Michael, J. (1975). Positive and negative reinforcement, a distinction that is no longer necessary; or a better way to talk about bad things. *Behaviorism, 3*, 33–44.

Michael, J. (1982). Distinguishing between the discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior, 37*, 149–155.

Michael, J. (1984). Verbal behavior. *Journal of the Experimental Analysis of Behavior, 42*, 363–376.

Michael, J. (1985). Two kinds of verbal behavior plus a possible third. *The Analysis of Verbal Behavior, 3*, 1–4. doi.org/10.1901/jeab.1984.42-363

Michael, J. (1986). Repertoire-altering effects of remote contingencies. *The Analysis of Verbal Behavior, 4*, 10–18.

Michael, J. (1987). Symposium on the experimental analysis of human behavior: Comments by the discussant. *Psychological Record, 37*, 37–42.

Michael, J. (1988). Establishing operations and the mand. *The Analysis of Verbal Behavior, 6*, 3–9.

Michael, J. (1991). *Verbal behavior: Objectives, exams, and exam answers.* Kalamazoo, MI: Western Michigan University.

Michael, J. (1992). Introduction I. In *Verbal behavior* by B. F. Skinner (Reprinted edition). Cambridge, MA: B. F. Skinner Foundation.

Michael, J. (1993). *Concepts and principles of behavior analysis*. Kalamazoo, MI: Society for the Advancement of Behavior Analysis.

Michael, J. (1993). Establishing operations. *The Behavior Analyst, 16*, 191–206.

Michael, J. (1995). What every student in behavior analysis ought to know: A system for classifying the multiple effects of behavioral variables. *The Behavior Analyst, 18*, 273–284.

Michael, J. (2000). Implications and refinements of the establishing operation concept. *Journal of Applied Behavior Analysis, 33*, 401–410.

Michael, J. (2004). *Concepts and principles of behavior analysis* (rev. ed.). Kalamazoo, MI: Society for the Advancement of Behavior Analysis.

Michael, J. (2006). Comment on Baron and Galizio. *The Behavior Analyst, 29*, 117–119.

Michael, J., & Malott, R. W. (2003). Michael and Malott's dialog on linguistic productivity. *The Analysis of Verbal Behavior, 19*, 115–118.

Michael, J., & Shafer, E. (1995). State notation for teaching about behavioral procedures. *The Behavior Analyst, 18*, 123–140.

Michael, J., Palmer, D. C., & Sundberg, M. L. (2011). The multiple control of verbal behavior. *The Analysis of Verbal Behavior, 27*, 3–22.

Michael, J. L. (2004). *Concepts and principles of behavior analysis* (rev. ed.). Kalamazoo, MI: Society for the Advancement of Behavior Analysis.

Michael, R. L., & Bernstein, D. J. (1991). The transient effect of acquisition history on generalization in a matching-to-sample task. *Journal of the Experimental Analysis of Behavior, 56*, 155–166.

Miguel, C. F. (2013). Jack Michael's motivation. *The Analysis of Verbal Behavior, 29*, 3–11.

Miguel, C. F. (2016). Common and intraverbal bidirectional naming. *The Analysis of Verbal Behavior, 32*, 125–138. doi:10.1007/s40616-016-0066-2

Miguel, C. F. (2017). The generalization of mands. *The Analysis of Verbal Behavior, 33*, 191–204.

Miguel, C. F. (2018). Problem-solving, bidirectional naming, and the development of verbal repertoires. *Behavior Analysis: Research and Practice*. Advance online publication. doi.org/10.1037/bar0000110

Miguel, C. F., & Kobari-Wright, V. V. (2013). The effects of tact training on the emergence of categorization and listener behavior in children with autism. *Journal of Applied Behavior Analysis, 46*, 669–773. doi:10.1002/jaba.62

Miguel, C. F., & Petursdottir, A. I. (2009). Naming and frames of coordination. In R. A. Rehfeldt & Y. Barnes-Holmes (Eds.), *Derived relational responding: Applications for learners with autism and other developmental disabilities* (pp. 129–148). Oakland, CA: New Harbinger Publications.

Miguel, C. F., Carr, J. E., & Michael, J. (2002). Effects of stimulus-stimulus pairing procedure on the vocal behavior of children diagnosed with autism. *The Analysis of Verbal Behavior, 18*, 3–13.

Miguel, C. F., Petursdottir, A. I., Carr, J. E., & Michael, J. (2008). The role of naming in stimulus categorization by preschool children. *Journal of the Experimental Analysis of Behavior, 89*, 383–405. doi:10.1901/jeab.2008-89-383

Miguel, C. F., Yang, H. G., Finn, H. E., & Ahern, W. H. (2009). Establishing derived textual control in activity schedules with children with autism. *Journal of Applied Behavior Analysis, 42*, 703–709.

Millenson, J. R. (1967). *Principles of behavioral analysis*. New York: Macmillan.

Miller, A. D., Hall, S. W., & Heward, W. L. (1995). Effects of sequential 1-minute time trials with and without inter-trial feedback and self-correction on general and special education students' fluency with math facts. *Journal of Behavioral Education, 5*, 319–345.

Miller, D. L., & Kelley, M. L. (1994). The use of goal setting and contingency contracting for improving children's homework performance. *Journal of Applied Behavior Analysis, 27*, 73–84.

Miller, D. L., & Stark, L. J. (1994). Contingency contracting for improving adherence in pediatric populations, *Journal of the American Medical Association, 271*(1), 81–83.

Miller, J. R., Lerman, D. C., & Fritz, J. N. (2010). An experimental analysis of negative reinforcement contingencies for adult-delivered reprimands. *Journal of Applied Behavior Analysis, 43*, 769–773.

Miltenberger, R. G. (2001). *Behavior modification: Principles and procedures* (2nd ed.). Belmont, CA: Wadsworth/Thomson Learning.

Miltenberger, R. G. (2008). Teaching safety skills to children: Prevention of firearm injury as an exemplar of best practice in assessment, training, and generalization of safety skills. *Behavior Analysis in Practice, 1*(1), 30–36. doi:10.1007/BF03391718

Miltenberger, R. G. (2016). *Behavior modification: Principles and procedures* (6th ed.). Boston: Cengage Learning.

Miltenberger, R. G., & Crosland, K. A. (2014). Parenting. In F. K. McSweeney & E. S. Murphy (Eds.), *The Wiley-Blackwell handbook of operant and classical conditioning* (pp. 509–531). Oxford, UK: John Wiley & Sons.

Miltenberger, R. G., & Fuqua, R. W. (1981). Overcorrection: A review and critical analysis. *The Behavior Analyst, 4*, 123–141.

Miltenberger, R. G., Flessner, C., Gatheridge, B., Johnson, B., Satterlund, M., & Egemo, K. (2004). Evaluation of behavior skills training to prevent gun play in children. *Journal of Applied Behavior Analysis, 37*, 513–516.

Miltenberger, R. G., Fuqua, R. W., & Woods, D. W. (1998). Applying behavior analysis with clinical problems: Review and analysis of habit reversal. *Journal of Applied Behavior Analysis, 31*, 447–469.

Miltenberger, R. G., Gatheridge, B. J., Satterlund, M., Egemo-Helm, K. R., Johnson, B. M., Jostad, C., & Flessner, C. A. (2005). Teaching safety skills to children to prevent gun play: An evaluation of in situ training. *Journal of Applied Behavior Analysis, 38*, 395–398. doi:10.1901/jaba.2005.130-04

Miltenberger, R. G., Rapp, J., & Long, E. (1999). A low-tech method for conducting real time recording. *Journal of Applied Behavior Analysis, 32*, 119–120.

Mineka, S. (1975). Some new perspectives on conditioned hunger. *Journal of Experimental Psychology: Animal Behavior Processes*, *104*, 143–148.

Mischel, H. N., Ebbesen, E. B., & Zeiss, A. R. (1972). Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology*, *16*, 204–218.

Mitchem, K. J., Young, K. R., West, R. P., & Benyo, J. (2001). CWPASM: A classwide peer-assisted self-management program for general education classrooms. *Education and Treatment of Children*, *24*, 3–14.

Moher, C. A., Gould, D. D., Hegg, E., & Mahoney, A. M. (2008). Non-generalized and generalized conditioned reinforcers: Establishment and validation. *Behavioral Interventions*, *23*, 13–38.

Montgomery, J., Storey, K., Post, M., & Lemley, J. (2011). The use of auditory prompting systems for increasing independent performance of students with autism in employment training. *International Journal of Rehabilitation Research*, *34*, 330–335.

Montoya-Rodríguez, M. M., Molina, F. J., & McHugh, L. (2017). A review of Relational Frame Theory research into deictic relational responding. *The Psychological Record*, *67*, 569–579.

Moore, J. (2000). Thinking about thinking and feeling about feeling. *The Behavior Analyst*, *23*(1), 45–56.

Moore, J. (2003). Behavior analysis, mentalism, and the path to social justice. *The Behavior Analyst*, *26*, 181–193.

Moore, J. (2008). *Conceptual foundations of radical behaviorism*. Cornwall-on-Hudson, NY: Sloan.

Moore, J. (2010). Behaviorism and the stages of scientific activity. *The Behavior Analyst*, *33*, 47–63.

Moore, J. (2011). A review of Baum's review of *Conceptual Foundations of Radical Behaviorism. Journal of the Experimental Analysis of Behavior*, *95*, 127–140.

Moore, J. (2015). *From a behavioral point of view: A psychological primer*. Cornwall-on-Hudson, NY: Sloan.

Moore, J., & Cooper, J. O. (2003). Some proposed relations among the domains of behavior analysis. *The Behavior Analyst*, *26*, 69–84.

Moore, J., & Shook, G. L. (2001). Certification, accreditation, and quality control in behavior analysis. *The Behavior Analyst*, *24*, 45–55.

Morales v. Turman, 364 F. Supp. 166 (E. D. Tx. 1973).

Moran, L., Stewart, I., McElwee, J., & Ming, S. (2010). Brief report: The training and assessment of relational precursors and abilities (TARPA): A preliminary analysis. *Journal of Autism and Developmental Disorders*, *40*, 1149–1153.

Moran, L., Stewart, I., McElwee, J., & Ming, S. (2014). Relational ability and language performance in children with autism spectrum disorders and typically developing children: A further test of the TARPA protocol. *The Psychological Record*, *64*, 233–251.

Moran, L., Walsh, L., Stewart, I., McElwee, J., & Ming, S. (2015). Correlating derived relational responding with linguistic and cognitive ability in children with Autism Spectrum Disorders. *Research in Autism Spectrum Disorders*, *19*, 32–43.

Morford, Z., Witts, B. N., Killingsworth, K. J., & Alavosius, M. P. (2014). Gamification: The intersection between behavior analysis and game design technologies. *The Behavior Analyst*, *37*, 25–40.

Morgan, D., Young, K. R., & Goldstein, S. (1983). Teaching behaviorally disordered students to increase teacher attention and praise in mainstreamed classrooms. *Behavioral Disorders*, *8*, 265–273.

Morgan, Q. E. (1978). *Comparison of two "Good Behavior Game" group contingencies on the spelling accuracy of fourth-grade students*. Unpublished master's thesis, The Ohio State University, Columbus, OH.

Morris, E. K. (1991). Deconstructing "technological to a fault." *Journal of Applied Behavior Analysis*, *24*, 411–416.

Morris, E. K. (2008). Sidney W. Bijou: The Illinois years. *The Behavior Analyst*, *31*, 179–203.

Morris, E. K. (2009). A case study in the misrepresentation of applied behavior analysis inautism: The Gernsbacher lectures. *The Behavior Analyst*, *32*, 205–230.

Morris, E. K. (2009). Sidney W. Bijou (1908–2009). *Inside Behavior Analysis*, *1*(1), 17–18.

Morris, E. K. (2013). The legacy of John B. Watson's behaviorist manifesto for applied behavior analysis. *Mexican Journal of Behavior Analysis*, *39*, 155–179.

Morris, E. K., & Smith, N. G. (2004). On the origin and preservation of cumulative record in its struggle for life as a favored term. *Journal of the Experimental Analysis of Behavior, 82*, 357–373.

Morris, E. K., Altus, A. E., & Smith, N. G. (2005). B. F. Skinner's contributions to applied behavior analysis. *The Behavior Analyst*, *28*, 99–131.

Morris, E. K., Hursh, D. E., Winston, A. S., Gelfand, D. M., Hartmann, D. P., Reese, H. W., & Baer, D. M. (1982). Behavior analysis and developmental psychology. *Human Development*, *25*, 340–364.

Morris, E. K., Todd, J. T., Midgley, B. D., Schneider, S. M., & Johnson, L. M. (1990). The history of behavior analysis: Some historiography and a bibliography. *The Behavior Analyst*, *13*, 131–158.

Morris, R. J. (1985). *Behavior modification with exceptional children: Principles and practices*. Glenview, IL: Scott, Foresman.

Morrison, R. S., Sainato, D. M., Benchaaban, D., & Endo, S. (2002). Increasing play skills of children with autism using activity schedules and correspondence training. *Journal of Early Intervention*, *25*, 58–72.

Morrow, J. (2017). Learning about John B. Watson. The story of a detour. *Operants, Quarter 1*, 6–7.

Morse, W. H., & Kelleher, R. T. (1977). Determinants of reinforcement and punishment. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 174–200). Upper Saddle River, NJ: Prentice Hall.

Morton, W. L., Heward, W. L., & Alber, S. R. (1998). When to self-correct? A comparison of two procedures on spelling performance. *Journal of Behavioral Education*, *8*, 321–335.

Mountjoy, P. T., & Cone, D. M. (1997). Another new journal? *The Psychological Record*: Volumes 1–V; 1937–1945. *The Psychological Record*, *47*, 3–20.

Mowrer, O. H. (1947). On the dual nature of learning—a reinterpretation of conditioning and problem solving. *Harvard Educational Review*, *17*, 102–148.

Mowrer, O. H. (1950). *Learning theory and personality dynamics*. New York: The Ronald Press Company.

Moxley, R. A. (1998). Treatment-only designs and student self-recording strategies for public school teachers. *Education and Treatment of Children*, *21*, 37–61.

Moxley, R. A. (2004). Pragmatic selectionism: The philosophy of behavior analysis. *The Behavior Analyst Today*, *5*, 108–125.

Moxon, P. D., Keenan, M., & Hine, L. (1993). Gender-role stereotyping and stimulus equivalence. *The Psychological Record*, *43*, 381–393.

Mrachko, A. A., Kostewicz, D. E., & Martin, W. P. (2017). Increasing positive and decreasing negative teacher responses to student behavior through training and feedback. *Behavior Analysis: Research and Practice*, *17*(3), 250–265. doi.org/10.1037/bar0000082

Mruzek, D. W., Cohen, C., & Smith, T. (2007). Contingency contracting with students with autism spectrum disorders in a public school setting. *Journal of Developmental and Physical Disabilities*, *19*, 103–114.

Mudford, O. C. (1995). Review of the gentle teaching data. *American Journal on Mental Retardation*, *99*, 345–355.

Mudford, O. C., & Cullen, C. (2016). Auditory integration training. In R. M. Foxx & J. A. Mulick (Eds.), *Controversial therapies for autism and intellectual disabilities: Fads, fashion, and science in professional practice* (2nd ed.) (pp. 136–153). New York: Routledge.

Mudford, O. C., Martin, N. T., Hui, J. K. Y., & Taylor, S. A. (2009). Assessing observer accuracy in continuous recording of rate and duration: Three algorithms compared. *Journal of Applied Behavior Analysis*, *42*, 527–539.

Mudford, O. C., Taylor, S. A., & Martin, N. T. (2009). Continuous recording and interobserver agreement algorithms reported in the *Journal of Applied Behavior Analysis* (1995–2005). *Journal of Applied Behavior Analysis*, *42*, 165–169.

Mudford, O. C., Zeleny, J. R., Fisher, W. W., Klum, M. E., & Owen, T. M. (2011). Calibration of observational measurement of rate of responding. *Journal of Applied Behavior Analysis*, *44*, 571–586. doi:10.1901/jaba.2011.44-571

Mueller, M. M., Moore, J. W., Doggett, R. A., & Tingstrom, D. H. (2000). The effectiveness of contingency-specific prompts in controlling bathroom graffiti. *Journal of Applied Behavior Analysis*, *33*, 89–92.

Mueller, M. M., Olmi, D. J., & Saunders, K. J. (2000). Recombinative generalization of within-syllable units in prereading children. *Journal of Applied Behavior Analysis*, *33*, 515–531.

Mulhern, T., Stewart, I., & McElwee, J. (2017). Investigating relational framing of categorization in young children. *The Psychological Record*, *67*, 519–536.

Mullins-Nelson, J. L., Salekin, R. T., & Leistico, A. M. R. (2006). Psychopathy, empathy, and perspective-taking ability in a community sample: Implications for the successful psychopathy concept. *International Journal of Forensic Mental Health*, *5*, 133–149.

Murphy, C., & Barnes-Holmes, D. (2009). Derived more-less relational mands in children diagnosed with autism. *Journal of Applied Behavior Analysis*, *42*, 253–268.

Murphy, C., Barnes-Holmes, D., & Barnes-Holmes, Y. (2005). Derived manding in children with autism: Synthesizing Skinner's Verbal Behavior with relational frame theory. *Journal of Applied Behavior Analysis*, *38*, 445–462.

Murphy, E. S., & Lupfer, G. J. (2011). *Basic principles of operant conditioning*. In F. K. McSweeney & E. S. Murphy (Eds.), *Wiley-Blackwell handbook of operant and classical conditioning* (pp. 167–194). Oxford: Wiley-Blackwell.

Murphy, G. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Murphy, R. J., Ruprecht, M. J., Baggio, P., & Nunes, D. L. (1979). The use of mild punishment in combination with reinforcement of alternate behaviors to reduce the self-injurious behavior of a profoundly retarded individual. *AAESPH Review*, *4*, 187–195.

Musser, E. H., Bray, M. A., Kehle, T. J., & Jenson, W. R. (2001). Reducing disruptive behaviors in students with serious emotional disturbance. *School Psychology Review*, *30*(2), 294–304.

Myer, J. S. (1971). Some effects of noncontingent aversive stimulation. In R. F. Brush (Ed.), *Aversive conditioning and learning* (pp. 469–536). NY: Academic Press.

Myers, S. M., Johnson, C. P., & the American Academy of Pediatrics Council on Children with Disabilities (2007). Management of children with autism spectrum disorders. *Pediatrics*, *120*, 1162–1182. doi:10.1542/peds.2007-2362. PMID 17967921

Myles, B. S., Moran, M. R., Ormsbee, C. K., & Downing, J. A. (1992). Guidelines for establishing and maintaining token economies. *Intervention in School and Clinic*, *27*(3), 164–169.

Najdowski, A. C., Chilingaryan, V., Berstrom, R., Granpeesheh, D., Balasanyan, S., Aguilar, B., & Tarbox, J. (2009). Comparison of data collection methods in a behavioral intervention program for children with pervasive developmental disorders: A replication. *Journal of Applied Behavior Analysis*, *42*, 827–832.

Najdowski, A. C., Wallace, M. D., Ellsworth, C. L., MacAleese, A. N., & Cleveland, J. M. (2008). Functional analyses and treatment of precursor behavior. *Journal of Applied Behavior Analysis*, *41*, 97–105. doi:10.1901/jaba.2008.41-97

Nakano, Y. (2004). Toward the establishment of behavioral ethics: Ethical principles of behavior analysis in the era of empirically supported treatment (EST). *Japanese Journal of Behavior Analysis*, *19*(1), 18–51.

Napolitano, D. A., Smith, T., Zarcone, J. R., Goodkin, & McAdam, D. B. (2010). Increasing response diversity in children with autism. *Journal of Applied Behavior Analysis*, *43*, 265–271.

National Aeronautics and Space Administration (NASA). (2019). *2018 fourth warmest year in continued warming trend, according to NASA, NOAA*. Washington, DC. Retrieved February 14, 2019, from https://www.nasa.gov/press-release/2018-fourth-warmest-year-in-continued-warming-trend-according-to-nasa-noaa

National Autism Center. (2015). *Findings and conclusions: National standards project, phase 2*. Randolf, MA: Author.

National Guardianship Association. (2016). Guardian. In *Guardianship terminology*. Bellefonte, PA: Author. Retrieved December 10, 2016, from guardianship.org/pdf/terminology.pdf

National Institute of Mental Health. (2008). *Autism Spectrum disorders: Pervasive developmental disorders*. NIH Publication no. 08-5511.

National Oceanic and Atmospheric Administration (NOAA). (2019). *2018 was 4th hottest year on record for the globe*. Washington, DC. Retrieved February 14, 2019, from https://www.noaa.gov/news/2018-was-4th-hottest-year-on-record-for-globe

National Reading Panel. (2000). Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups. (NIH Pub No. 00-4754). Bethesda, MD: National Institute of Child Health and Human Development. [Available at https://www.nichd.nih.gov/publications/pubs/nrp/Pages/report.aspx].

Navarro, J. I., Aguilar, M., Aguilar, C., Alcalde, C., & Marchena, E. (2007). Positive behavioral intervention in children who were wards of the court attending a mainstream school. *Psychological Reports*, *101*, 1067–1078.

Neef, N. A., Bicard, D. F., & Endo, S. (2001). Assessment of impulsivity and the development of self-control in students with attention-deficit/hyperactivity disorder. *Journal of Applied Behavior Analysis*, *34*, 397–408. doi:10.1901/jaba.2001.34-397

Neef, N. A., Bicard, D. F., Endo, S., Coury, D. L., & Aman, M. G. (2005). Evaluation of pharmacological treatment of impulsivity by students with attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis*, *38*, 135–146.

Neef, N. A., Iwata, B. A., & Page, T. J. (1980). The effects of interspersal training versus high density reinforcement on spelling acquisition and retention. *Journal of Applied Behavior Analysis*, *13*, 153–158.

Neef, N. A., Iwata, B. A., & Page, T. J. (1986). Ethical standards in behavioral research. In A. Poling & R. W. Fuqua (Eds.), *Research methods in applied behavior analysis* (pp. 233–263). New York: Springer.

Neef, N. A., Lensbower, J., Hockersmith, I., DePalma, V., & Gray, K. (1990). In vivo versus stimulation training: An interactional analysis of range and type of training exemplars. *Journal of Applied Behavior Analysis*, *23*, 447–458.

Neef, N. A., Mace, F. C., & Shade, D. (1993). Impulsivity in students with serious emotional disturbance: The interactive effects of reinforcer rate, delay, and quality. *Journal of Applied Behavior Analysis*, *26*, 37–52.

Neef, N. A., Mace, F. C., Shea, M. C., & Shade, D. (1992). Effects of reinforcer rate and reinforcer quality on time allocation: Extensions of matching theory to educational settings. *Journal of Applied Behavior Analysis*, *25*, 691–699.

Neef, N. A., Marckel, J., Ferreri, S. J., Bicard, D. F., Endo, S., Aman, M. G., . . . Armstrong, N. (2005). Behavioral assessment of impulsivity: A comparison of children with and without attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis*, *38*, 23–37. doi:10.1901/jaba.2005.146-02

Neef, N. A., Parrish, J. M., Hannigan, K. F., Page, T. J., & Iwata, B. A. (1990). Teaching self-catheterization skills to children with neurogenic bladder complications. *Journal of Applied Behavior Analysis*, *22*, 237–243.

Neidert, P. L., Iwata, B. A., & Dozier, C. L. (2005). Treatment of multiply controlled problem behavior with procedural variations of differential reinforcement. *Exceptionality*, *13*, 45–53.

Neuringer, A. (2003). Reinforced variability and creativity. In K. A. Lattal & P. N. Chase (Eds.), *Behavior theory and philosophy* (pp. 323–338). New York: Kluwer Academic/Plenum.

Neuringer, A. (2004). Reinforced variability in animals and people: Implications for adaptive action. *American Psychologist*, *59*, 891–906.

Neuringer, A. (2009). Operant variability and the power of reinforcement. *The Behavior Analyst Today*, *10*, 319–343.

Neuringer, A., & Jensen, G. (2013). Operant variability. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. 513–546). Washington, DC: American Psychological Association.

Nevin, J. A. (1992). An integrative model for the study of behavioral momentum. *Journal of the Experimental Analysis of Behavior*, *57*, 301–316.

Nevin, J. A. (1998). Choice and momentum. In W. O'Donohue (Ed.), *Learning and behavior therapy* (pp. 230–251). Boston: Allyn and Bacon.

Nevin, J. A. (2009). Stimuli, reinforcers, and the persistence of behavior. *The Behavior Analyst*, *32*, 285–291.

Nevin, J. A., & Grace, R. C. (2000). Behavioral momentum and the law of effect. *Behavioral and Brain Sciences*, *23*, 73–130.

Nevin, J. A., Mandell, C., & Atak, J. R. (1983). The analysis of behavioral momentum. *Journal of the Experimental Analysis of Behavior*, *39*, 49–59.

Nevin, J. A., Tota, M. E., Torquato, R. D., & Shull, R. L. (1990). Alternative reinforcement increases resistance to change: Pavlovian or operant contingencies? *Journal of the Experimental Analysis of Behavior*, *53*, 359–379.

New York State Department of Health Early Intervention Program (1999). *Clinical practice guideline: Report of the recommendations, autism/pervasive developmental disorders, assessment and intervention for young children*. Publication #4215. Health Education Services, P.O. Box 7126, Albany, NY 12224, http://www.health.state.ny.us/nysdoh/eip/menu.htm

Newhouse-Oisten, M. K., Peck, K. M., Conway, A. A., & Frieder, J. E. (2017). Ethical considerations for interdisciplinary collaboration with prescribing professionals. *Behavior Analysis in Practice*, *10*, 145–153. Published online 05 Apr 2017. doi 10.1007/s40617-017-0184-x.

Newman, R. S., & Golding, L. (1990). Children's reluctance to seek help with school work. *Journal of Educational Psychology*, *82*, 92–100.

Newstrom, J., McLaughlin, T. F., & Sweeney, W. J. (1999). The effects of contingency contracting to improve the mechanics of written language with a middle school student with behavior disorders. *Child & Family Behavior Therapy*, *21*(1), 39–48.

Nichols, T. (2017). *The death of expertise: The campaign against established knowledge and why it matters*. New York: Oxford University Press.

Nikopoulos, C. K., & Keenan, M. (2004). Effects of video modeling on social initiations by children with autism. *Journal of Applied Behavior Analysis*, *37*, 93–96.

Ninness, C., Barnes-Holmes, D., Rumph, R., McCuller, G., Ford, A. M., Payne, R., . . . Elliott, M. P. (2006). Transformations of mathematical and stimulus functions. *Journal of Applied Behavior Analysis*, *39*, 299–321.

Ninness, C., Dixon, M., Barnes-Holmes, D., Rehfeldt, R. A., Rumph, R., McCuller, G., . . . McGinty, J. (2009). Constructing and deriving reciprocal trigonometric relations: A functional analytic approach. *Journal of Applied Behavior Analysis*, *42*, 191–208.

Ninness, C., Rumph, R., McCuller, G., Harrison, C., Ford, A. M., & Ninness, S. (2005). A functional analytic approach to computer-interactive mathematics. *Journal of Applied Behavior Analysis*, *38*, 1–22.

Ninness, H. A. C., Fuerst, J., & Rutherford, R. D. (1991). Effects of self-management training and reinforcement on the transfer of improved conduct in the absence of supervision. *Journal of Applied Behavior Analysis*, *24*, 499–508.

Niwayama, K., & Tanaka-Matsumi, J. (2016). Effects of teachers' self-recording of behavior-specific praise on children's academic engagement in the classroom. *Japanese Journal of Educational Psychology*, *64*, 598–609.

Noell, G. H., Call, N. A., & Ardoin, S. P. (2011). Building complex repertoires from discrete behaviors by establishing stimulus control, behavior chains, and strategic behavior. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 250–269). New York: Guilford Press.

Noell, G. H., VanDerHeyden, A. M., Gatti, S. L., & Whitmarsh, E. L. (2001). Functional assessment of the effects of escape and attention on students' compliance during instruction. *School Psychology Quarterly*, *16*, 253–269.

Nolan, J. D. (1968). Self-control procedures in the modification of smoking behavior. *Journal of Consulting and Clinical Psychology*, *32*, 92–93.

Noldus, L. P. P. J., Trienes, R. J. H., Hendriksen, H. J., & Jansen, R. G. (2000). The observer video-pro: New software for the collection, management, and presentation of time-structured data from videotapes and digital media files. *Behavior Research Methods*, *32*, 197–206.

Norman, J. E. (1977). *The effects of programmed instructional materials: Parent training in contingency contracting* (Doctoral dissertation). Retrieved from ProQuest Dissertations Publishing, 7805898.

Normand, M. P. (2008). Science, skepticism, and applied behavior analysis. *Behavior Analysis in Practice*, *1*(2), 42–49.

Normand, M. P. (2009). Much ado about nothing? Some comments on Skinner's definition of verbal behavior. *The Behavior Analyst*, *32*, 185–190.

Normand, M. P., Kestner, K., & Jessel, J. (2010). An analysis of stimuli that influence compliance during the high-probability instruction sequence. *Journal of Applied Behavior Analysis*, 43, 735–738. doi:10.1901/jaba.2010.43-735

North, S. T., & Iwata, B. A. (2005). Motivational influences on performance maintained by food reinforcement. *Journal of Applied Behavior Analysis*, *38*, 317–333.

Northeastern University. (1983). *Barnabus: A study in complex behavior chaining*. Boston: Northeastern University Film Project.

Northup, J. (2000). Further evaluation of the accuracy of reinforcer surveys: A systematic replication. *Journal of Applied Behavior Analysis*, *33*, 335–338.

Northup, J., George, T., Jones, K., Broussard, C., & Vollmer, T. R. (1996). A comparison of reinforcer assessment methods: The utility of verbal and pictorial choice procedures. *Journal of Applied Behavior Analysis*, *29*, 201–212.

Northup, J., Wacker, D., Sasso, G., Steege, M., Cigrand, K., Cook, J., & DeRaad, A. (1991). A brief functional analysis of aggressive and alternative behavior in an outclinic setting. *Journal of Applied Behavior Analysis*, *24*, 509–522. doi:10.1901/jaba.1991.24-509

Nosik, M. R., & Carr, J. E. (2015). On the distinction between the motivating operation and setting event concepts. *The Behavior Analyst*, *38*, 219–223.

Novak, G., & Pelaez, M. (2004). *Child and adolescent development: A behavioral systems approach*. Thousand Oaks, CA: Sage Publications.

Nuzzolo-Gomez, R., & Greer, R. D. (2004). Emergence of untaught mands or tacts of novel adjective-object pairs as a function of instructional history. *The Analysis of Verbal Behavior*, *20*, 63–76.

O'Brien, F. (1968). Sequential contrast effects with human subjects. *Journal of the Experimental Analysis of Behavior*, *11*, 537–542.

O'Brien, S., & Karsh, K. G. (1990). Treatment acceptability, consumer, therapist, and society. In A. C. Repp & N. N. Singh (Eds.), *Perspectives on the use of nonaversive and aversive interventions for persons with developmental disabilities* (pp. 503–516). Sycamore, IL: Sycamore.

O'Conner, R. T., Lerman, D. C., & Fritz, J. N. (2010). Effects of number and locations of bins on plastic recycling at a university. *Journal of Applied Behavior Analysis*, *43*, 711–715

O'Donnell, J. (2001). The discriminative stimulus for punishment or S^{Dp}. *The Behavior Analyst*, *24*, 261–262.

O'Donohue, W., & Ferguson, K. E. (2011). Behavior analysis and ethics. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 489–497). New York, London: Guilford Press.

O'Hora, D., Peláez, M., & Barnes-Holmes, D. (2005). Derived relational responding and performance on verbal subtests of the WAIS-III. *The Psychological Record*, *55*, 155–175.

O'Hora, D., Peláez, M., Barnes-Holmes, D., Rae, G., Robinson, K., & Chaudhary, T. (2008). Temporal relations and intelligence: Correlating relational performance with performance on the WAIS-III. *The Psychological Record*, *58*, 569–584.

O'Leary, K. D. (1977). Teaching self-management skills to children. In D. Upper (Ed.), *Perspectives in behavior therapy*. Kalamazoo, MI: Behaviordelia.

O'Leary, K. D., & O'Leary, S. G. (Eds.). (1972). *Classroom management: The successful use of behavior modification*. New York: Pergamon.

O'Leary, K. D., Kaufman, K. F., Kass, R. E., & Drabman, R. S. (1970). The effects of loud and soft reprimands on the behavior of disruptive students. *Exceptional Children*, *37*, 145–155.

O'Leary, P. N., Miller, M. M., Olive, M. L., & Kelly, A. N. (2017). Blurred lines: Ethical implications of social media for behavior analysts. *Behavior Analysis in Practice*, *10*, 45–51. Published online 06 Jan 2015. doi 10.1007/s40617-014-0033-0

O'Leary, S. G., & Dubey, D. R. Applications of self-control procedures by children: A review. *Journal of Applied Behavior Analysis*, *12*, 449–465.

O'Neill, R. E., Bundock, K., Kladis, K., & Hawken, L. S. (2015). Acceptability of functional behavioral assessment procedures to special educators and school psychologists. *Behavioral Disorders*, *41*, 51–66.

O'Neill, R. E., Horner, R. H., Albin, R. W., Sprague, J. R., Storey, K., & Newton, J. S. (1997). *Functional assessment for problem behavior: A practical handbook* (2nd ed.). Pacific Grove, CA: Brooks/Cole.

O'Reilly, M. F. (1995). Functional analysis and treatment of escape-maintained aggression correlated with sleep deprivation. *Journal of Applied Behavior Analysis*, *28*, 225–226.

O'Reilly, M. F. (1997). Functional analysis of episodic self-injury correlated with recurrent otitis media. *Journal of Applied Behavior Analysis*, *30*, 165–167.

O'Reilly, M. F., Aguilar, J., Fragale, C., Lang, R., Edrisinha, C., Sigafoos, J., . . . Didden, R. (2012). Effects of motivating operation manipulation on the maintenance of mands. *Journal of Applied Behavior Analysis*, *45*, 443–447.

O'Reilly, M. F., Lancioni, G. E., King, L., Lally, G., & Dhomhnaill, O. N. (2000). Using brief assessments to evaluate aberrant behavior maintained by attention. *Journal of Applied Behavior Analysis*, *33*, 109–112.

O'Reilly, M. F., Lang, R., Davis, T., Rispoli, M., Machalicek, W., Sigafoos, J., et al. (2009). A systematic examination of different parameters of presession exposure to tangible stimuli that maintain problem behavior. *Journal of Applied Behavior Analysis*, *42*, 773–783.

O'Reilly, M. F., Sigafoos, J., Lancioni, G., Rispoli, M., Lang, R., Chan, J., . . . Langthorne, P. (2007). Manipulating the behavior-altering effect of the motivating operation: Examination of the influence on challenging behavior during leisure activities. *Research in Developmental Disabilities*, *29*, 333–340.

O'Sullivan, J. L. (1999). Adult guardianship and alternatives. In R. D. Dinerstein, S. S. Herr, & J. L. O'Sullivan (Eds.), *A guide to consent* (pp. 7–37). Washington, DC: American Association on Intellectual and Developmental Disabilities.

Odom, S. L., Hoyson, M., Jamieson, B., & Strain, P. S. (1985). Increasing handicapped preschoolers' peer social interactions: Cross-setting and component analysis. *Journal of Applied Behavior Analysis, 18*, 3–16.

Odum, A. L. (2011). Delay discounting: I'm a k, you're a k. *Journal of the Experimental Analysis of Behavior, 96*, 427–439.

Okinaka, T., & Shimazaki, T. (2011). The effects of prompting and reinforcement on safe behavior of bicycle and motorcycle riders. *Journal of Applied Behavior Analysis, 44*, 671–674.

Olaff, H. S., Ona, H. N., & Holth, P. (2017, January 23). Establishment of naming in children with autism through multiple response-exemplar training. *Behavioral Development Bulletin, 22*(1), 67–85.

Oleson, C. R., & Baker, J. C. (2014). Teaching mands to older adults with dementia. *The Analysis of Verbal Behavior, 30*(2), 113–127.

Oliver, C. O., Oxener, G., Hearn, M., & Hall, S. (2001). Effects of social proximity on multiple aggressive behaviors. *Journal of Applied Behavior Analysis, 34*, 85–88.

Ollendick, T., Matson, J., Esveldt-Dawson, K., & Shapiro, E. (1980). An initial investigation into the parameters of overcorrection. *Psychological Reports, 39*, 1139–1142.

Ollendick, T. H., Matson, J. L., Esvelt Dawson, K., & Shapiro, E. S. (1980). Increasing spelling achievement: An analysis of treatment procedures utilizing an alternating treatments design. *Journal of Applied Behavior Analysis, 13*, 645–654.

Olympia, D. W., Sheridan, S. M., Jenson, W. R., & Andrews, D. (1994). Using student-managed interventions to increase homework completion and accuracy. *Journal of Applied Behavior Analysis, 27*, 85–99.

Operant Systems Inc. (2017). *ABA Data NotePad—Skills*. New York: Operant Systems Inc. Retrieved December 27, 2017, at https://itunes.apple.com/us/app/aba-data-notepad-skills/id1248744125?mt=8

Ortiz, K. R., & Carr, J. E. (2000). Multiple-stimulus preference assessments: A comparison of free-operant and restricted-operant formats. *Behavioral Interventions, 15*, 345–353.

Osborne, J. G. (1969). Free-time as a reinforcer in the management of classroom behavior. *Journal of Applied Behavior Analysis, 2*, 113–118.

Osborne, J. G., & Gatch, M. B. (1989). Stimulus equivalence and receptive reading by hearing-impaired preschool chidren. *Language, Speech, and Hearing Services in Schools, 20*, 63–75.

Osnes, P. G., & Lieblein, T. (2003). An explicit technology of generalization. *The Behavior Analyst Today, 3*, 364–374.

Osnes, P. G., Guevremont, D. C., & Stokes, T. F. (1984). If I say I'll talk more, then I will: Correspondence training to increase peer-directed talk by socially withdrawn children. *Behavior Modification, 10*, 287–299.

Østvik, L., Eikeseth, S., & Klintwall, L. (2012). Grammatical constructions in typical developing children: Effects of explicit reinforcement, automatic reinforcement and parity. *The Analysis of Verbal Behavior, 28*, 73–82. doi.org/10.1007/BF03393108

Ottenbacher, K. J. (1993). Interrater agreement of visual analysis in single-subject decisions: Quantitative review and analysis. *American Journal on Mental Retardation, 98*, 135–142.

Owens, R. E., Jr. (2012). *Language development: An introduction* (8th ed.). Upper Saddle River, NJ: Pearson.

Pace, G. A., Iwata, B. A., Cowdery, G. E., Andree, P. J., & McIntyre, T. (1993). Stimulus (instructional) fading during extinction of self-injurious escape behavior. *Journal of Applied Behavior Analysis, 26*, 205–212.

Pace, G. M., & Troyer, E. A. (2000). The effects of a vitamin supplement on the pica of a child with severe mental retardation. *Journal of Applied Behavior Analysis, 33*, 619–622.

Pace, G. M., Ivancic, M. T., Edwards, G. L., Iwata, B. A., & Page, T. J. (1985). Assessment of stimulus preference and reinforcer value with profoundly retarded individuals. *Journal of Applied Behavior Analysis, 18*, 249–255.

Paclawskyj, T., Matson, J., Rush, K., Smalls, Y., & Vollmer, T. (2000). Questions about behavioral function (QABF): Behavioral checklist for functional assessment of aberrant behavior. *Research in Developmental Disabilities, 21*, 223–229.

Paclawskyj, T. R., & Vollmer, T. R. (1995). Reinforcer assessment for children with developmental disabilities and visual impairments. *Journal of Applied Behavior Analysis, 28*, 219–224.

Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes, 11*, 429–452.

Page, T. J., & Iwata, B. A. (1986). Interobserver agreement: History, theory, and current methods. In A. Poling & R. W. Fuqua (Eds.), *Research methods in applied behavior analysis* (pp. 92–126). New York: Plenum Press.

Page, T. J., Finney, J. W., Parrish, J. M., & Iwata, B. A. (1983). Assessment and reduction of food stealing in Prader-Willi children. *Applied Research in Mental Retardation, 4*(3), 219–228.

Pagoto, S., Waring, M. E., May, C. N., Ding, E. Y., Kunz, W. H., Hayes, R., & Oleski, J. L. (2016). Adapting behavioral interventions for social media delivery. *Journal of Medical Internet Research, 18*(1) e24. doi:10.2196/jmir.5086

Paliliunas, D., Belisle, J., Barron, B., & Dixon, M. R. (in press). Early hierarchical multiply-controlled relations evident in a game of "I spy" with children with autism. *Behavior Analysis in Practice*.

Palmer, D. C. (1991). A behavioral interpretation of memory. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 261–279). Reno, NV: Context Press.

Palmer, D. C. (1996). Achieving parity: The role of automatic reinforcement. *Journal of the Experimental Analysis of Behavior, 65*, 289–290.

Palmer, D. C. (1998). On Skinner's rejection of S-R psychology. *The Behavior Analyst, 21*, 93–96.

Palmer, D. C. (1998). The speaker as listener: The interpretation of structural regularities in verbal behavior. *The Analysis of Verbal Behavior, 15*, 3–16.

Palmer, D. C. (2004). Data in search of a principle: A review of *Relational frame theory: A post-Skinnerian account of human language and cognition*. *Journal of the Experimental Analysis of Behavior, 81*, 189–204. doi:10.1901/jeab.2004.81-189

Palmer, D. C. (2006). Joint control: A discussion of recent research. *The Analysis of Verbal Behavior, 22*, 209–215.

Palmer, D. C. (2011). Consideration of private events is required in a comprehensive science of behavior. *The Behavior Analyst, 34*, 201–207.

Palmer, D. C. (2012). The role of atomic repertoires in the interpretation of complex behavior. *The Behavior Analyst, 35*, 59–73.

Palmer, D. C. (2016). On intraverbal control and the definition of the intraverbal. *The Analysis of Verbal Behavior, 33*, 107–124. doi 10.1007/s40616-017-0086-6

Palmer, D. C. (2018). A behavioral interpretation of aesthetics. *The Psychological Record*. https://doi.org/10.1007/s40732-018-0306-z

Panyan, M., Boozer, H., & Morris, N. (1970). Feedback to attendants as a reinforcer for applying operant techniques. *Journal of Applied Behavior Analysis, 3*, 1–4.

Parker, L. H., Cataldo, M. F., Bourland, G., Emurian, C. S., Corbin, R. J., & Page, J. M. (1984). Operant treatment of orofacial dysfunction in neuromuscular disorders. *Journal of Applied Behavior Analysis, 17*, 413–427.

Parra, I., & Ruiz, F. J. (2016). The effect on intelligence quotient of training fluency in relational frames of coordination. *International Journal of Psychology and Psychological Therapy, 16*, 1–12.

Parsons, M. B., Reid, D. H., Bentley, E., Inman, A., & Lattimore, L. P. (2012). Identifying indices of happiness and unhappiness among adults with autism: Potential targets for behavioral assessment and intervention. *Behavior Analysis in Practice, 5*, 15–25.

Parsons, M. B., Reid, D. H., Reynolds, J., & Bumgarner, M. (1990). Effects of chosen versus assigned jobs on the work performance of persons with severe handicaps. *Journal of Applied Behavior Analysis, 23*, 253–258.

Partington, J. W. (2006). *The assessment of basic language and learning skills, revised. (The ABLLS-R)*. Pleasant Hill, CA: Behavior Analysts, Inc.

Partington, J. W., & Sundberg, M. L. (1998). *The assessment of basic language and learning skills (The ABLLS)*. Pleasant Hill, CA: Behavior Analysts, Inc.

Patel, M. R., Piazza, C. C., Kelly, M. L., Ochsner, C. A., & Santana, C. M. (2001). Using a fading procedure to increase fluid consumption in a child with feeding problems. *Journal of Applied Behavior Analysis, 34*, 357–360.

Patel, M. R., Piazza, C. C., Kelly, M. L., Ochsner, C. A., & Santana, C. M. (2001). Using a fading procedure to increase fluid consumption in a child with feeding problems. *Journal of Applied Behavior Analysis, 34*, 357–360.

Patel, M. R., Reed, G. K., Piazza, C. C., Bachmeyer, M. H., Layer, S. A., & Pabico, R. S. (2006). An evaluation of the high-probability instructional sequence to increase acceptance of food and decrease inappropriate behavior in children with pediatric feeding disorders. *Research in Developmental Disabilities, 27*, 430–442.

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia.

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial boys. Vol. 4: A social interactional approach*. Eugene, OR: Castalia.

Pauwels, A. A., Ahearn, W. H., & Cohen, S. J. (2015). Recombinative generalization of tacts through matrix training with individuals with autism spectrum disorder. *The Analysis of Verbal Behavior*, *31*, 200–214. doi:10.1007/s40616-015-0038-y

Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex* (W. H. Grant, Trans.). London, UK: Oxford University Press.

Payne, S. W., & Dozier, C. L. (2013). Positive reinforcement as treatment for problem behavior maintained by negative reinforcement. *Journal of Applied Behavior Analysis*, *46*, 699–703.

Pelaez-Nogueras, M., Gewirtz, J. L., Field, T., Cigales, M., Malphurs, J., Clasky, S., & Sanchez, A. (1996). Infants' preference for touch stimulation in face-to-face interactions. *Journal of Applied Developmental Psychology*, *17*, 199–213.

Peleg, R., Martin, N., & Holth, P. (2017). Variability and resistance to change of different response positions in response sequences emitted under continuous and lag reinforcement schedules. *European Journal of Behavior Analysis*, *18*, 264–275.

Pelios, L., Morren, J., Tesch, D., & Axelrod, S. (1999). The impact of functional analysis methodology on treatment choice for self-injurious and aggressive behavior. *Journal of Applied Behavior Analysis*, *32*, 185–195.

Pena, T., Pitts, R. C., & Galizio, M. (2006). Identity matching-to-sample with olfactory stimuli in rats. *Journal of the Experimental Analysis of Behavior*, *85*, 203–221.

Pence, S. T., & St. Peter, C. C. (2015). Evaluation of of treatment-integrity errors on mand acquisition. *Journal of Applied Behavior Analysis*, *48*, 575–589. doi:10.1002/jaba.238

Penney, H., Prirchard, D., Hoerger, M., & Mace, F. C. (2016). Conditioned suppression of compulsive behavior in an adolescent diagnosed with obsessive-compulsive disorder. *European Journal of Behavior Analysis*, *18*, 192–199.

Pennington, B., & McComas, J. J. (2017). Effects of the Good Behavior Game across classroom contexts. *Journal of Applied Behavior Analysis*, *50*, 176–180. doi:10.1002/jaba.357

Pennypacker, H. S. (1981). On behavioral analysis. *The Behavior Analyst*, *3*, 159–161.

Pennypacker, H. S. (1994). A selectionist view of the future of behavior analysis in education. In R. Gardner III, D. M. Sainato, J. O. Cooper, T. E. Heron, W. L. Heward, J. Eshleman, & T. A. Grossi (Eds.), *Behavior analysis in education: Focus on measurably superior instruction* (pp. 11–18). Monterey, CA: Brooks/Cole.

Pennypacker, H. S., Gutierrez, A., & Lindsley, O. R. (2003). *Handbook of the Standard Celeration Chart*. Gainesville, FL: Xerographics.

Penrod, B., Gardella, L., & Fernand, J. (2012). An evaluation of a progressive high-probability instructional sequence combined with low-probability demand fading in the treatment of food selectivity. *Journal of Applied Behavior Analysis*, *45*, 527–537. doi:10.1901/jaba.2012.45-527

Pérez-González, L. A., & García-Asenjo, L. (2016). Emergence of intraverbals with antonyms derived from relations with verbal and nonverbal stimuli. *The Psychological Record*, *66*, 351–368. doi.org/10.1007/s40732-016-0177-0

Pérez-González, L. A., & Moreno-Sierra, V. (1999). Equivalence class formation in elderly persons. *Psicothema*, *11*, 325–336.

Pérez-González, L. A., García-Asenjo, L., Williams, G., & Carnerero, J. J. (2007). Emergence of intraverbal antonyms in children with pervasive developmental disorder. *Journal of Applied Behavior Analysis*, *40*, 697–701.

Pérez-González, L. A., Herszlikowicz, K., & Williams, G. (2008). Stimulus relations analysis and the emergence of novel intraverbals. *The Psychological Record*, *58*, 95–129.

Pérez-González, L. A., Pastor, A., & Carnerero, J. J. (2014). Observing tacting increases uninstructed tacts in children with autism. *The Analysis of Verbal Behavior*, *30*, 62–68. doi.org/10.1007/s40616-013-0003-6

Pérez-González, L. A., Salamah, J., & García-Asenjo, L. (2018). Emergence of intraverbals with categories as responses after learning intraverbals with elements in reverse stimulus response relations. *European Journal of Behavior Analysis*, online. doi:10.1080/15021149.2018.1465755

Perez, W. F., & de Rose, J. C. (2010). Recombinative generalization: An exploratory study in musical reading. *The Analysis of Verbal Behavior*, *26*, 51–55.

Perone, M., & Galizio, M. (1987). Variable-interval schedules of timeout from avoidance. *Journal of the Experimental Analysis of Behavior*, *47*, 97–113.

Perone, M., & Hursh, D. E. (2013). Single-case experimental designs. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. 107–126). Washington, DC: American Psychological Association.

Perrin, C. J., Dayton, E., Quigley, J., Prater, A., & Landis, J. D. (2018). Further evaluation of latency measures in the assessment and treatment of severe self-injurious behavior. *Behavioral Interventions*, *33*(2), 196–204. doi:10.1002/bin.1520

Perrin, C. J., Fredrick, A., & Klick, S. (2016). Effects of public posting and goals setting on team performance in a residential setting. *Behavioral Interventions*, *31*, 300–308. doi:10.1002/bin.1451

Peters, M., & Heron, T. E. (1993). When the best is not good enough: An examination of best practice. *Journal of Special Education*, *26*(4), 371–385.

Peterson, C., & Bossio, L. M. (1998). Learned helplessness. In R. C. Curtis (Ed.), *Self-defeating behaviors: Experimental research, clinical impressions, and practical implications* (pp. 235–257). New York: Plenum.

Peterson, I., Homer, A. L., & Wonderlich, S. A. (1982). The integrity of independent variables in behavior analysis. *Journal of Applied Behavior Analysis*, *15*, 477–492.

Peterson, K. M., Volkert, V. M., & Zeleny, J. R. (2015). Increasing self-drinking for children with feeding disorders. *Journal of Applied Behavior Analysis*, *48*, 436–441.

Peterson, N., & Ledoux, S. F. (2014). *An introduction to verbal behavior* (2nd ed.). Canton, NY: ABC of Canton.

Peterson, R. F. (1968). Some experiments on the organization of a class of imitative behavior. *Journal of Applied Behavior Analysis*, *1*, 225–235.

Petscher, E. S., Rey, C., & Bailey, J. A. (2009). A review of empirical support for differential reinforcement of alternative behavior. *Research in Developmental Disabilities*, *30*, 409–425.

Petursdottir, A. I. (2018). The current status of the experimental analysis of verbal behavior. *Behavior Analysis Research and Practice*, *18*, 151–168.

Petursdottir, A. I., & Aguilar, G. (2016). Order of stimulus presentation influences children's acquisition in receptive identification tasks. *Journal of Applied Behavior Analysis*, *49*, 58–68.

Petursdottir, A. I., & Carr, J. E. (2011). A review of recommendations for sequencing receptive and expressive language instruction. *Journal of Applied Behavior Analysis*, *46*, 859–876.

Petursdottir, A. I., & Devine, B. (2017). The impact of *Verbal Behavior* on the scholarly literature from 2005 to 2016. *The Analysis of Verbal Behavior*, *33*, 212–228. doi.org/10.1007/s40616-017-0089-3

Petursdottir, A. I., & Haflidadóttir, R. D. (2009). A comparison of four strategies for teaching a small foreign-language vocabulary. *Journal of Applied Behavior Analysis*, *42*, 685–690. doi.org/10.1901/jaba.2009.42-685

Petursdottir, A. I., Carr, J. E., Lechago, S. A., & Almason, S. M. (2008). An evaluation of intraverbal training and listener training for teaching categorization skills. *Journal of Applied Behavior Analysis*, *41*, 53–68. doi.org/10.1901/jaba.2008.41-53

Petursdottir, A. I., Olafsdottir, A. R., & Aradottir, B. (2008). The effects of tact and listener training on the emergence of bidirectional intraverbal relations. *Journal of Applied Behavior Analysis*, *41*, 411–415.

Pfadt, A., & Wheeler, D. J. (1995). Using statistical process control to make data-based clinical decisions. *Journal of Applied Behavior Analysis*, *28*, 349–370.

Pfadt, A., Cohen, I. L., Sudhalter, V., Romanczyk, R. G., & Wheeler, D. J. (1992). Applying statistical process control to clinical data: An illustration. *Journal of Applied Behavior Analysis*, *25*, 551–560.

Phillips, C. L., Iannaccone, J. A., Rooker, G. W., & Hagopian, L. P. (2017). Noncontingent reinforcement for the treatment of severe problem behavior: An analysis of 27 consecutive applications. *Journal of the Experimental Analysis of Behavior*, *50*, 357–376.

Phillips, E. L., Phillips, E. A., Fixen, D. L., & Wolf, M. M. (1971). Achievement Place: Modification of the behaviors of predelinquent boys with a token economy. *Journal of Applied Behavior Analysis*, *4*, 45–59.

Piazza, C. C., & Fisher, W. (1991). A faded bedtime with response cost protocol for treatment of multiple-sleep problems in children. *Journal of Applied Behavior Analysis*, *24*, 129–140.

Piazza, C. C., Bowman, L. G., Contrucci, S. A., Delia, M. D., Adelinis, J. D., & Goh, H.-L. (1999). An evaluation of the properties of attention as

reinforcement for destructive and appropriate behavior. *Journal of Applied Behavior Analysis*, *32*, 437–449.

Piazza, C. C., Fisher, W. W., Hagopian, L. P., Bowman, L. G., & Toole, L. (1996). Using a choice assessment to predict reinforcer effectiveness. *Journal of Applied Behavior Analysis*, *29*, 1–9.

Piazza, C. C., Fisher W. W., Hanley, G. P., Remick, M. L., Contrucci, S. A., & Aitken, T. L. (1997). The use of positive and negative reinforcement in the treatment of escape-maintained destructive behavior. *Journal of Applied Behavior Analysis*, *30*, 279–298.

Piazza, C. C., Fisher, W. W., & Sherer, M. (1997). Treatment of multiple sleep disorders in children with developmental disabilities. *Developmental Medicine and Child Neurology*, *39*, 414–418.

Piazza, C. C., Moses, D. R., & Fisher, W. W. (1996). Differential reinforcement of alternative behavior and demand fading in the treatment of escape maintained destructive behavior. *Journal of Applied Behavior Analysis*, *29*, 569–572.

Piazza, C. C., Patel, M. R., Gulotta, C. S., Sevin, B. M., & Layer, S. A. (2003). On the relative contributions of positive reinforcement and escape extinction in the treatment of food refusal. *Journal of Applied Behavior Analysis*, *36*, 309–324.

Piazza, C. C., Roane, H. S., Kenney, K. M., Boney, B. R., & Abt, K. A. (2002). Varying response effort in the treatment of pica maintained by automatic reinforcement. *Journal of Applied Behavior Analysis*, *35*, 233–246.

Pierce, J. M., Spriggs, A. D., Gast, D. L., & Luscre, D. (2013). Effects of visual activity schedules on independent classroom transitions for students with autism. *International Journal of Disability Development and Education*, *60*(3), 253–269.

Pierce, W. D., & Epling, W. F. (1995). *Behavior analysis and learning*. Englewood Cliffs, NJ: Prentice-Hall.

Pierrel, R., & Sherman, J. G. (1963). Train your pet the Barnabus way. *Brown Alumni Monthly*, February 1963, pp. 8–14.

Pietras, C. J., & Hackenberg, T. D. (2005). Response-cost punishment via token-loss with pigeons. *Behavioural Processes*, *69*, 343–356.

Pilgrim, C. (2015). Stimulus control and generalization. In F. D. DiGennaro Reed & D. D. Reed (Eds.), *Autism service delivery: Bridging the gap between science and practice* (pp. 25–74). New York: Springer Science.

Pilgrim, C., & Galizio, M. (1990). Relations between baseline contingencies and equivalence probe performances. *Journal of the Experimental Analysis of Behavior*, *54*, 213–224.

Pilgrim, C., Boye, J., Hogan, W. G., & Groff, R. V. (in press). Equivalence-class formation and three-term contingencies. *Journal of the Experimental Analysis of Behavior*.

Pilgrim, C., Chambers, L., & Galizio, M. (1995). Reversal of baseline relations and stimulus equivalence: II: Children. *Journal of the Experimental Analysis of Behavior*, *63*, 239–254.

Pilgrim, C., Jackson, J., & Galizio, M. (2000). Acquisition of arbitrary conditional discriminations by young typically developing children. *Journal of the Experimental Analysis of Behavior*, *73*, 177–193.

Pinkelman, S. E., & Horner, R. H. (2016). Improving implementation of function-based interventions: Self-monitoring, data collections, and data review. *Journal of Positive Behavior Interventions*, *18*, 1–11. doi:10.1177/1098300716683634

Pitts, L., & Dymond, S. (2012). Increasing compliance of children with autism: Effects of programmed reinforcement for high-probability requests and varied inter-instruction intervals. *Research in Autism Spectrum Disorders*, *6*, 135–143. doi:10.1016/j.rasd.2011.03.013

Plavnick, J. B., Ferreri, S. J., & Maupin, A. N. (2010). The effects of self-monitoring on the procedural integrity of a behavioral intervention for young children with developmental disabilities. *Journal of Applied Behavior Analysis*, *43*, 315–320.

Ploog, B. O. (2010). Stimulus overselectivity four decades later: A review of the literature and its implications for current research in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *40*, 1332–1349. doi:10.1007/s10803-010-0990-2

Poche, C., Brouwer, R., & Swearingen, M. (1981). Teaching self-protection to young children. *Journal of Applied Behavior Analysis*, *14*, 169–176.

Poche, C., McCubbrey, H., & Munn, T. (1982). The development of correct toothbrushing technique in preschool children. *Journal of Applied Behavior Analysis*, *15*, 315–320. doi.org/10.1901/jaba.1982.15-315

Poche, C., Yoder, P., & Miltenberger, R. (1988). Teaching self-protection skills to children using television techniques. *Journal of Applied Behavior Analysis*, *21*, 253–261.

Poling, A. (2010). Looking to the future: Will behavior analysis survive and prosper? *The Behavior Analyst*, *33*, 7–17.

Poling, A. (2010). Progressive-ratio schedules and applied behavior analysis. *Journal of Applied Behavior Analysis*, *43*, 347–349.

Poling, A., & Normand, M. (1999). Noncontingent reinforcement: An inappropriate description of time-based schedules that reduce behavior. *Journal of Applied Behavior Analysis*, *32*, 237–238.

Poling, A., & Ryan, C. (1982). Differential-reinforcement-of-other-behavior schedules: Therapeutic applications. *Behavior Modification*, *6*, 3–21.

Poling, A., Lotfizadeh, A., & Edwards, T. L. (2017). Predicting reinforcement: Utility of the motivating operations concept. *The Behavior Analyst*, *40*, 49–56.

Poling, A., Methot, L. L., & LeSage, M. G. (Eds.). (1995). *Fundamentals of behavior analytic research*. New York: Plenum Press.

Poling, A., Valverde, E., Beyene, N., Mulder, C., Cox, C., Mgode, G., & Edwards, T. L. (2017). Active tuberculosis detection by pouched rats in 2014: More than 2,000 new patients found in two countries. *Journal of Applied Behavior Analysis*, *50*, 165–169.

Poling, A., Weetjens, B. J., Cox, C., Beyene, N., Bach, H., & Sully, A. (2010). Teaching giant African pouched rats to find landmines: Operant conditioning with real consequences. *Behavior Analysis in Practice*, *3*(2), 19–25.

Poling, A., Weetjens, B., Cox, C. Beyene, N. W. Bach, H., & Sully A. (2011a). Using trained pouched rats to detect land mines: Another victory for operant conditioning. *Journal of Applied Behavior Analysis*, *44*, 351–355.

Poling, A., Weetjens, B., Cox, C., Beyene, N., Durgin, A., & Mahoney, A. (2011b). Tuberculosis detection by giant African pouched rats. *The Behavior Analyst*, *34*(1), 47–54.

Polson, D. A., & Parsons, J. A. (2000). Selection-based versus topography-based responding: An important distinction for stimulus equivalence. *The Analysis of Verbal Behavior*, *17*, 105–128.

Popkin, J., & Skinner, C. H. (2003). Enhancing academic performance in a classroom serving students with serious emotional disturbance: Interdependent group contingencies with randomly selected components. *School Psychology Review*, *32*, 282–295.

Poulson, C. L. (1983). Differential reinforcement of other-than-vocalization as a control procedure in the conditioning of infant vocalization rate. *Journal of Experimental Child Psychology*, *36*, 471–489.

Poulson, C. L. (2003). Generalized imitation: An empirical and conceptual analysis. In K. S. Budd & T. F. (Eds.). *A small matter of proof: The legacy of Donald M. Baer* (pp. 101–112). Reno, NV: Context Press.

Poulson, C. L, Kymissis, E., Reeve, K. F., Andreatos, M., & Reeve, L. (1991). Generalized vocal imitation in infants. *Journal of Experimental Child Psychology*, *51*, 267–279.

Powell, J., & Azrin, N. (1968). Behavioral engineering: Postural control by a portable operant apparatus. *Journal of Applied Behavior Analysis*, *1*, 63–71.

Powell. R. W. (1969). The effect of reinforcement magnitude upon responding under fixed-ratio schedules. *Journal of the Experimental Analysis of Behavior*, *12*, 605–608.

Powers, R. B., Osborne, J. G., & Anderson, E. G. (1973). Positive reinforcement of litter removal in the natural environment. *Journal of Applied Behavior Analysis*, *6*, 579–586.

Premack, D. (1959). Toward empirical behavioral laws: I. Positive reinforcement. *Psychological Review*, *66*, 219–233.

President's Council On Physical Fitness. (2016). https://www.presidentschallenge.org/challenge/physical/benchmarks.shtml

Pritchard, D., Penney, H., & Mace, R. C. (2018). The *ACHIEVE!* program: A point and level system for reducing severe problem behavior. *Behavioral Interventions*, *33*, 41–55. doi:10.1002/bin.1506

Pritchard, J. (2010). Virtual rewards for driving green. *The Behavior Analyst*, *33*, 185–187.

Progar, P. R., North, S. T., Bruce, S. S., Dinovi, B. J., Nan, P. A., Eberman, E. M., . . . Nussbaum, C. N. (2001). Putative behavioral history effects and aggression maintained by escape from therapists. *Journal of Applied Behavior Analysis*, *34*, 69–72.

Protopopova, A., Kisten, D., & Wynne, C. (2016). Evaluating a humane alternative to the bark collar: Automated differential reinforcement of not barking in a home-alone setting. *Journal of Applied Behavior Analysis*, *49*, 735–744.

Pryor, K. (1999). *Don't shoot the dog! The new art of teaching and training* (rev. ed.). New York: Bantam Books.

Pryor, K. (2002). *Click to win! Clicker training for the show ring*. Waltham, MA: Sunshine Books.

Pryor, K., & Chamberlain, L. (2016). *The modern principles of shaping*. Retrieved January 15, 2018, from https://www.clickertraining.com/files/the-modern-principles-of-shaping-SF-edits.pdf

Pryor, K., & Norris, K. S. (1991). *Dolphin societies: Discoveries and puzzles*. Berkeley, CA: University of California Press.

Pryor, K., & Ramirez, K. (2014). Modern animal training: A transformative technology. In F. K. McSweeney & E. S. Murphy (Eds.), *Wiley-Blackwell handbook of operant and classical conditioning* (pp. 455–482). Oxford: Wiley-Blackwell.

Pryor, K., Haag, R., & O'Reilly, J. (1969). The creative porpoise: Training for novel behavior. *Journal of the Experimental Analysis of Behavior*, *12*, 653–661.

Pryor, M. (2005). *Click Air: Clicker trained flight instruction, Part 1*. Retrieved August 6, 2016, from http://www.clickertraining.com/node/385.

Putnam, B. C., & Tiger, J. H. (2016). Assessing generative braille responding following training in a matching-to-sample format. *Journal of Applied Behavior Analysis*, *49*, 751–767.

Pytte, C. L., & Fienup, D. M. (2012). Using equivalence-based instruction to increase efficiency in teaching neuroanatomy. *Journal of Undergraduate Neuroscience Education*, *10*, A125–A131.

Qi, C. H., Barton, E. E., Collier, M., & Lin, Y. (2017). A systematic review of single-case research studies on using video modeling interventions to improve social communication skills for individuals with autism spectrum disorder. *Focus on Autism and Other Developmental Disabilities*. doi.org/10.1177/1088357617741282

Quinn, M., Miltenberger, R., Abreu, A., & Narozanick, T. (2017). An evaluation of public posting and graphical feedback to enhance the performance of competitive dancers. *Behavior Analysis in Practice*, *10*, 1–11.

Quinn, M. J., Miltenberger, R. G., & Fogel, V. A. (2015). Using TAGTeach to improve the proficiency of dance movements. *Journal of Applied Behavior Analysis*, *48*, 11–14.

Rachlin, H. (1995). *Self-control*. Cambridge. MA: Harvard University Press.

Rachlin, H. (2016). Self-control based on soft-commitment. *The Behavior Analyst*, *39*, 259–268.

Rachlin, H. (2018). Where do we go from here? *Behavior Analysis: Research and Practice*, *18*(2), 119–123.

Rachlin, H., & Green, L. (1972). Commitment, choice, and self-control. *Journal of the Experimental Analysis of Behavior*, *17*, 15–22.

Rader, L., Sidener, T. M., Reeve, K. F., Sidener, D. W., Delmolino, L., Miliotis, A., & Carbone, V. J. (2014). Stimulus-stimulus pairing of vocalizations: A systematic replication. *The Analysis of Verbal Behavior*, *30*, 69–74.

Radley, K. C., Dart, E. H., Battaglia, A. A., & Ford, W. B. (2018). A comparison of two procedures for assessing preference in a classroom setting. *Behavior Analysis in Practice*, doi.org/10.1007/s40617-018-0244-x

Rafferty, L. A. (2010). Step-by-step: Teaching students to self-monitor. *Teaching Exceptional Children*, *43*(2), 50–58.

Raiff, B. R., & Dallery, J. (2010). Internet-based contingency management to improve adherence with blood glucose testing recommendations for teens with Type 1 diabetes. *Journal of Applied Behavior Analysis 43*, 487–491. doi:10.1901/jaba.2010.43-487

Raiff, B. R., Barry, V. B., Ridenour, T. A., & Jitnarin, N. (2016). Internet-based incentives increase blood glucose testing with a non-adherent, diverse sample of teens with type 1 diabetes mellitus: A randomized controlled trial. *Translational Behavioral Medicine: Practice, Policy and Research*, *6*(2), 179–188. doi:10.1007/s13142-016-0397-5

Raiff, B. R., Jarvis, B. P., & Dallery, J. (2016). Text-message reminders plus incentives increase adherence to antidiabetic medication in adults with Type 2 diabetes. *Journal of Applied Behavior Analysis*, *49*, 947–953. doi:10.1002/jaba.337

Ramon, D., Yu, C. T., Martin, G. L., & Martin, T. (2015). Evaluation of a self-instructional manual to teach multiple-stimulus without replacement preference assessments. *Journal of Behavioral Education*, *24*, 289–303. doi 10.1007/s10864-015-9222-3

Randell, T., & Remington, B. (1999). Equivalence relations between visual stimuli: The functional role of naming. *Journal of the Experimental Analysis of Behavior*, *71*, 395–415.

Randell, T., & Remington, B. (2006). Equivalence relations, contextual control, and naming. *Journal of the Experimental Analysis of Behavior*, *86*, 337–354. doi:10.1901/jeab.2006.82-05

Randell, T., & Remington, B. (2013). Equivalence relations, contextual control and naming. *Journal of the Experimental Analysis of Behavior*, *86*, 337–354.

Rantz, W. G., & Van Houten, R. (2013). A feedback intervention to increase digital and paper checklist performance in technically advanced aircraft simulation. *Journal of Applied Behavior Analysis*, *44*, 145–150.

Rapp, J. T., Carroll, R. A., Stangeland, L., Swanson, G., & Higgins, W. J. (2011). A comparison of reliability measures for continuous and discontinuous recording methods: Inflated agreement scores with partial interval recording and momentary time sampling for duration events. *Behavior Modification*, *35*, 389–402. http://dx.doi.org/10.1177/0145445511405512

Rapp, J. T., Colby-Dirksen, A. M., Michalski, D. N., Carroll, R. A., & Lindenberg, A. M. (2008). Detecting changes in simulated events using partial-interval recording and momentary time sampling. *Behavioral Interventions*, *23*, 237–269.

Rapp, J. T., Cook, J. L., McHugh, C., & Mann, K.R. (2017). Decreasing stereotypy using NCR and DRO with functionally matched stimulation: Effects on targeted and nontargeted stereotypy. *Behavior Modification*, *41*, 45–83.

Rapp, J. T., Miltenberger, R. G., Galensky, T. L., Ellingson, S. A., & Long, E. S. (1999). A functional analysis of hair pulling. *Journal of Applied Behavior Analysis*, *32*, 329–337.

Rasey, H. W., & Iversen, I. H. (1993). An experimental acquisition of maladaptive behaviors by shaping. *Journal of Behavior Therapy and Experimental Psychiatry*, *24*, 37–43.

Readdick, C. A., & Chapman, P. L. (2000). Young children's perceptions of time out. *Journal of Research in Childhood Education*, *15*(1), 81–87.

Redditor Jasonp55. (2018). *Correlation between autism diagnosis and organic food sales*. Retrieved August 14, 2018, from https://boingboing.net/2013/01/01/correlation-between-autism-dia.html

Reed, D. D., DiGennaro Reed, F. D., Jenkins, S. R., & Hirst, J. M. (2014). The zeitgeist of behavior analytic research in the 21st century: A keyword analysis. *The Behavior Analyst Today*, *14*, 17–25. doi:10.1037/h0101282

Reese, E. P. (1966). *The analysis of human operant behavior*. Dubuque, IA: Wm. C. Brown Company Publishers.

Reese, H. W. (1968). *The perception of stimulus relations: Discrimination learning and transposition*. New York: Academic Press.

Reese, H. W. (2005a). A conceptual analysis of selectionism: Part I. *Behavioral Development Bulletin*, *12*(1), 8–12.

Reese, H. W. (2005b). A conceptual analysis of selectionism: Part II. *Behavioral Development Bulletin*, *12*(1), 13–16.

Rehfeldt, R. A. (2003). Establishing contextual control over generalized equivalence relations. *The Psychological Record*, *53*, 415–428.

Rehfeldt, R. A. (2011). Toward a technology of derived stimulus relations: An analysis of articles published in the *Journal of Applied Behavior Analysis*, 1992–2009. *Journal of Applied Behavior Analysis*, *44*, 109–119.

Rehfeldt, R. A., & Barnes-Holmes, Y. (Eds.). (2009). *Derived relational responding: Applications for learners with autism and other developmental disabilities: A progressive guide to change*. New Oakland, CA: Harbinger Publications.

Rehfeldt, R. A., & Chambers, M. C. (2003). Functional analysis and treatment of verbal perseverations displayed by an adult with autism. *Journal of Applied Behavior Analysis*, *36*, 259–261.

Rehfeldt, R. A., & Root, S. L. (2005). Establishing derived requesting skills in adults with severe developmental disabilities. *Journal of Applied Behavior Analysis*, *38*, 101–105.

Rehfeldt, R. A., Dillen, J. E., Ziomek, M. M., & Kowalchuk, R. K. (2007). Assessing relational learning deficits in perspective-taking in children with high-functioning autism spectrum disorder. *The Psychological Record*, *57*, 23–47.

Reich, W. T. (1988). Experimental ethics as a foundation for dialogue between health communications and health-care ethics. *Journal of Applied Communications Research*, *16*, 16–28.

Reid, D. H., Parsons, M. B., Green, C. W., & Browning, L. B. (2001). Increasing one aspect of self-determination among adults with severe multiple disabilities in supported work. *Journal of Applied Behavioral Analysis*, *34*, 341–344.

Reid, D. H., Parsons, M. B., & Jensen, J. M. (2017). Maintaining staff performance following a training intervention: Suggestions from a 30-year case example. *Behavior Analysis in Practice*, *10*, 12–21. doi:10.1007/s40617-015-0101-0

Reid, D. H., Parsons, M. B., Phillips, J. F., & Green, C. W. (1993). Reduction of self-injurious hand mouthing using response blocking. *Journal of Applied Behavior Analysis*, *26*, 139–140.

Reid, R. (1996). Research in self-monitoring with students with learning disabilities: The present, the prospects, the pitfalls. *Journal of Learning Disabilities*, *29*, 317–331.

Reid, R., & Harris, K. R. (1993). Self-monitoring attention versus self-monitoring of performance: Effects on attention and academic performance. *Exceptional Children*, *60*, 29–40.

Reid, R., & Nelson, J. R. (2002). The utility, acceptability, and practicality of functional behavioral assessment for students with high-incidence problem behaviors. *Remedial and Special Education*, *23*(1), 15. Retrieved from http://libproxy.library.wmich.edu/login?url=https://search-proquest-com.libproxy.library.wmich.edu/docview/236322700?accountid=15099

Reimers, T. M., & Wacker, D. P. (1988). Parents' ratings of the acceptability of behavior treatment recommendations made in an outpatient clinic: A preliminary analysis of the influence of treatment effectiveness. *Behavioral Disorders*, *14*, 7–15.

Reitman, D., & Drabman, R. S. (1999). Multifaceted uses of a simple timeout record in the treatment of a noncompliant 8-year-old boy. *Education and Treatment of Children*, *22*(2), 136–145.

Reitman, D., & Gross, A. M. (1996). Delayed outcomes and rule-governed behavior among "noncompliant" and "compliant" boys: A replication and extension. *The Analysis of Verbal Behavior*, *13*, 65–77.

Repp, A. C., & Horner, R. H. (Eds.). (1999). *Functional analysis of problem behavior: From effective assessment to effective support*. Belmont, CA: Wadsworth.

Repp, A. C., & Singh, N. N. (Eds.). (1990). *Perspectives on the use of nonaversive and aversive interventions for persons with developmental disabilities*. Sycamore, IL: Sycamore.

Repp, A. C., Barton, L. E., & Brulle, A. R. (1983). A comparison of two procedures for programming the differential reinforcement of other behaviors. *Journal of Applied Behavior Analysis*, *16*, 435–445.

Repp, A. C., Dietz, D. E. D., Boles, S. M., Dietz, S. M., & Repp, C. F. (1976). Differences among common methods for calculating interobserver agreement. *Journal of Applied Behavior Analysis*, *9*, 109–113.

Repp, A. C., Harman, M. L., Felce, D., Vanacker, R., & Karsh, K. L. (1989). Conducting behavioral assessments on computer collected data. *Behavioral Assessment*, *2*, 249–268.

Rescorla, R. A. (2004). Spontaneous recovery. *Learning and Memory*, *11*, 501–509. doi:10.1101/lm.77504

Reynolds, B., Gast, D. L., & Luscre, D. (2014). Self-management by kindergarten students with disabilities in the general education classroom. *Journal of Positive Behavior Interventions*, *16*(3), 137–148. doi:10.1177/1098300713483176

Reynolds, G. S. (1961). Behavioral contrast. *Journal of the Experimental Analysis of Behavior*, *4*, 57–71.

Reynolds, G. S. (1968). *A primer of operant conditioning*. Glenview, IL: Scott, Foresman.

Reynolds, G. S. (1975). *A primer of operant conditioning* (rev. ed.). Glenview, IL: Scott, Foresman and Company.

Reynolds, N. J., & Risley, T. R. (1968). The role of social and material reinforcers in increasing talking of a disadvantaged preschool child. *Journal of Applied Behavior Analysis*, *1*, 253–262.

Rhode, G., Morgan, D. P., & Young, K. R. (1983). Generalization and maintenance of treatment gains of behaviorally handicapped students from resource rooms to regular classrooms using self-evaluation procedures. *Journal of Applied Behavior Analysis*, *16*, 171–188.

Ribeiro, D. M., Elias, N. C., Goyos, C., & Miguel, C. F. (2010). The effects of listener training on the emergence of tact and mand signs by individuals with intellectual disabilities. *The Analysis of Verbal Behavior*, *26*, 65–72. doi.org/10.1007/ BF03393084

Richard, E. B., Taylor, B., DeQuinzio, J. A., & Katz, B. (2010). *A comparison of two data collection and graphing systems: Paper and pencil and TeachMe*. 36th Annual Convention of the Association for Behavior Analysis International, San Antonio, TX.

Richardson, A. R., Lerman, D. C., Nissen, M. A., Luck, K. M., Neal, A. E., Bao, S., & Tsami, L. (2017). Can pictures promote the acquisition of sight-word reading? An evaluation of two potential instructional strategies. *Journal of Applied Behavior Analysis*, *50*, 67–86. doi:10.1002/jaba.354

Ricketts, R. W., Goza, A. B., & Matese, M. (1993). A 4-year follow-up of treatment of self-injury. *Journal of Behavior Therapy and Experimental Psychiatry*, *24*(1), 57–62.

Rincover, A. (1981). *How to use sensory extinction*. Austin, TX: PRO-ED.

Rincover, A., & Koegel, R. L. (1975). Setting generality and stimulus control in autistic children. *Journal of Applied Behavior Analysis*, *8*, 235–246.

Rincover, A., & Newsom, C. D. (1985). The relative motivational properties of sensory and edible reinforcers in teaching autistic children. *Journal of Applied Behavior Analysis*, *18*, 237–248. doi:10.1901/jaba.1985.18-237

Rincover, A., Cook, R., Peoples, A., & Packard, D. (1979). Sensory extinction and sensory reinforcement principles for programming multiple adaptive behavior change. *Journal of Applied Behavior Analysis*, *12*, 221–233.

Rindfuss, J. B., Al-Attrash, M., Morrison, H., & Heward, W. L. (1998, May). *Using guided notes and response cards to improve quiz and exam scores in an eighth grade American history class*. Paper presented at 24th Annual Convention of the Association for Behavior Analysis, Orlando, FL.

Ring, B. M., Sigurdsson, S. O., Eubanks, S. L., & Silverman, K. (2014). Reduction of classroom noise levels using group contingencies *Journal of Applied Behavior Analysis*, *47*, 840–844.

Ringdahl, J. E., & St. Peter, C. (2017). Resurgence: The unintended maintenance of problem behavior. *Education and Treatment of Children*, *40*(1), 7–26.

Ringdahl, J. E., Kitsukawa, K., Andelman, M. S., Call, N., Winborn, L. C., Barretto, A., & Reed, G. K. (2002). Differential reinforcement with and without instructional fading. *Journal of Applied Behavior Analysis*, *35*, 291–294.

Ringdahl, J. E., Vollmer, T. R., Borrero, J. C., & Connell, J. E. (2001). Fixed-time schedule effects as a function of baseline reinforcement rate. *Journal of Applied Behavior Analysis*, *34*, 1–15.

Ringdahl, J. E., Vollmer, T. R., Marcus, B. A., & Roane, H. S. (1997). An analogue evaluation of environmental enrichment: The role of stimulus preference. *Journal of Applied Behavior Analysis*, *30*, 203–216. doi:10.1901/jaba.1997.30-203

Riordan, M. M., Iwata, B. A., Finney, J. W., Wohl, M. K., & Stanley, A. E. (1984). Behavioral assessment and treatment of chronic food refusal in handicapped children. *Journal of Applied Behavior Analysis*, *17*, 327–341.

Risley, T. (1996). Get a life! In L. K. Koegel, R. L. Koegel, & G. Dunlap (Eds.), *Positive behavioral support* (pp. 425–437). Baltimore: Brookes.

Risley, T. (2005). Montrose M. Wolf (1935–2004). *Journal of Applied Behavior Analysis*, *38*, 279–287.

Risley, T. R. (1969, April). *Behavior modification: An experimental-therapeutic endeavor*. Paper presented at the Banff International Conference on Behavior Modification, Banff, Alberta, Canada.

Risley, T. R., & Hart, B. (1968). Developing correspondence between the non-verbal and verbal behavior of preschool children. *Journal of Applied Behavior Analysis*, *1*, 267–281.

Risley, T. R., & Reynolds, N. J. (1970). Emphasis as a prompt for verbal imitation. *Journal of Applied Behavior Analysis*, *3*, 185–190. doi: 10.1901/jaba.1970.3-185

Rispoli, M., Burke, M. D., Hatton, H., Ninci, J., Zaini, S., & Sanchez, L. (2015). Training Head Start teachers to conduct trial-based functional analysis of challenging behavior. *Journal of Positive Behavior Interventions*, *17*, 235–244. doi:10.1177/1098300715577428

Riviere, V., Becquet, M., Peltret, E., Facon, B., & Darcheville, J. C. (2011). Increasing compliance with medical examination requests directed to children with autism: Effects of a high-probability request procedure. *Journal of Applied Behavior Analysis*, *44*, 193–197. doi:10.1901/jaba.2011.44-193

Roane, H. (2008). On the applied use of progressive-ratio schedules of reinforcement. *Journal of Applied Behavior Analysis*, *41*, 155–161.

Roane, H. S., Kelly, M. L., & Fisher, W. W. (2003). The effects of noncontingent access to food on the rate of object mouthing across three settings. *Journal of Applied Behavior Analysis*, *36*, 579–582.

Roane, H. S., Lerman, D. C., Kelley, M. E., & Van Camp, C. M. (1999). Within-session patterns of responding during functional analyses: The role of establishing operations in clarifying behavioral function. *Research in Developmental Disabilities*, *20*, 73–89.

Roane, H. S., Lerman, D. C., & Vorndran, C. M. (2001). Assessing reinforcers under progressive schedule requirements. *Journal of Applied Behavior Analysis*, *34*, 145–167.

Roane, H. S., Vollmer, T. R., Ringdahl, J. E., & Marcus, B. A. (1998). Evaluation of a brief stimulus preference assessment. *Journal of Applied Behavior Analysis*, *31*, 605–620.

Roberton, T., Daffern, M., & Bucks, R. S. (2012). Emotion regulation and aggression. *Aggression and Violent Behavior*, *17*, 72–82.

Roberts-Pennell, D., & Sigafoos, J. (1999). Teaching young children with developmental disabilities to request more play using the behavior chain interruption strategy. *Journal of Applied Research in Intellectual Disabilities*, *12*, 100–112.

Robin, A., Schneider, M., & Dolnick, M. (1976). The turtle technique: An extended case study of self-control in the classroom. *Psychology in the Schools*, *13*, 449–453.

Robinson, P. W., Newby, T. J., & Gansell, S. L. (1981). A token system for a class of underachieving hyperactive children. *Journal of Applied Behavior Analysis*, *14*, 307–315.

Rock, M. L. (2005). Use of strategic self-monitoring to enhance academic engagement, productivity, and accuracy of students with and without exceptionalities. *Journal of Positive Behavior Interventions*, *7*, 3–17.

Rodriguez, M. A. (2011). Cash is king—how OBM helped a North American telecommunications organization obtain $76 million in receivables. *Journal of Organizational Behavior Management*, *31*(3), 163–178. doi: 10.1080/01608061.2011.589733

Rodriguez, N. M., Thompson, R. H., Schlichenmeyer, K., & Stocco, C. S. (2012). Functional analysis and treatment of arranging and ordering by individuals with an autism spectrum disorder. *Journal of Applied Behavior Analysis*, *45*, 1–22.

Roediger, H. L. (2004, March). What happened to behaviorism? *APS Observer*, *17*, 40–42.

Rohn, D. (2002). Case study: Improving guitar skills. In R. W. Malott & H. Harrison, *I'll stop procrastinating when I get around to it: Plus other cool ways to succeed in school and life using behavior analysis to get your act together* (pp. 8–4). Kalamazoo, MI: Department of Psychology, Western Michigan University.

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales* (5th ed.). Itasca, IL: Riverside.

Rolider, A., & Van Houten, R. (1984). The effects of DRO alone and DRO plus reprimands on the undesirable behavior of three children in home settings. *Education and Treatment of Children*, *7*, 17–31.

Rolider, A., & Van Houten, R. (1984). Training parents to use extinction to eliminate nighttime crying by gradually increasing the criteria for ignoring crying. *Education and Treatment of Children*, *7*, 119–124.

Rolider, A., & Van Houten, R. (1985). Suppressing tantrum behavior in public places through the use of delayed punishment mediated by audio recordings. *Behavior Therapy*, *16*, 181–194.

Rolider, A., & Van Houten, R. (1993). The interpersonal treatment model. In R. Van Houten & S. Axelrod (Eds.), *Behavior analysis and treatment* (pp. 127–168). New York: Plenum Press.

Rolider, A., Williams, L., Cummings, A., & Van Houten, R. (1991). The use of a brief movement restriction procedure to eliminate severe inappropriate behavior. *Journal of Behavior Therapy and Experimental Psychiatry*, *22*, 23–30.

Roll, J. M. (2014). Behavioral pharmacology: A brief overview. In F. K. McSweeney & E. S. Murphy (Eds.), *The Wiley-Blackwell handbook of operant and classical conditioning* (pp. 611–625). Oxford, UK: John Wiley & Sons.

Roll, J. M., Madden, G. J., Rawson, R., & Petry, N. M. (2009). Facilitating the adoption of contingency management for the treatment of substance abuse disorders. *Behavior Analysis in Practice*, *2*(1), 4–13.

Romanczyk, R. G. (1977). Intermittent punishment of self-stimulation: Effectiveness during application and extinction. *Journal of Consulting and Clinical Psychology*, *45*, 53–60.

Romanczyk, R. G., & McEachin, J. (2016). *Comprehensive models of autism spectrum disorder treatment: Points of divergence and convergence*. New York: Springer.

Romaniuk, C., Miltenberger, R., Conyers, C., Jenner, N., Jurgens, M., & Ringenberg, C. (2002). The influence of activity choice on problem behaviors maintained by escape versus attention. *Journal of Applied Behavioral Analysis*, *35*, 349–362.

Romanowich, P., & Lamb, R. J. (2010). Effects of escalating and descending schedules of incentives on cigarette smoking in smokers without plans to quit. *Journal of Applied Behavior Analysis*, *43*, 357–367. doi:10.1901/jaba.2010.43-357

Romanowich, P., & Lamb, R. J. (2015). The effects of fixed versus escalating reinforcement schedules on smoking abstinence. *Journal of Applied Behavior Analysis*, *48*, 25–37. doi:10.1002/jaba.185

Romeo, F. F. (1998). The negative effects of using a group contingency system of classroom management. *Journal of Instructional Psychology*, *25*(2), 130–133.

Romer, L. T., Cullinan, T., & Schoenberg, B. (1994). General case training of requesting: A demonstration and analysis. *Education and Training in Mental Retardation*, *29*, 57–68.

Rosales, R., & Rehfeldt, R. A. (2007). Contriving transitive conditioned establishing operations to establish derived manding skills in adults with severe developmental disabilities. *Journal of Applied Behavior Analysis*, *40*, *105–121*.

Rosales, R., Rehfeldt, R. A., & Lovett, S. (2011). Effects of multiple exemplar training on the emergence of derived relations in preschool children learning a second language. *The Analysis of Verbal Behavior*, *27*, 61–74.

Rosales-Ruiz, J., & Baer, D. M. (1997). Behavioral cusps: A developmental and pragmatic concept for behavior analysis. *Journal of Applied Behavior Analysis*, *30*, 533–544.

Roscoe, E. M., & Fisher, W. W. (2008). Evaluation of an efficient method for training staff to implement stimulus preference assessments. *Journal of Applied Behavior Analysis*, *41*, 249–254. doi:10.1901/jaba.2008.41–249

Roscoe, E. M., Iwata, B. A., & Goh, H. L. (1998). A comparison of noncontingent reinforcement and sensory extinction as treatments for self-injurious behavior. *Journal of Applied Behavior Analysis*, *31*, 635–646. doi:10.1901/jaba.1998.31-635

Roscoe, E. M., Iwata, B. A., & Kahng, S. (1999). Relative versus absolute reinforcement effects: Implications for preference assessments. *Journal of Applied Behavior Analysis*, *32*, 479–493.

Roscoe, E. M., Rooker, G. W., Pence, S. T., & Longworth, L. J. (2009). Assessing the utility of a demand assessment for functional analysis. *Journal of Applied Behavior Analysis*, *42*, 819–825.

Rosenberg, N.E. & Schwartz, I.S. (2018). Guidance or compliance: What makes an ethical behavior analyst? Behavior Analysis Practice. https://doi.org/10.1007/s40617-018-00287-5

Rosenfarb, I. S., Newland, M. C., Brannon, S. E., & Howey, D. S. (1992). Effects of self-generated rules on the development of schedule-controlled behavior. *Journal of the Experimental Analysis of Behavior*, *58*, 107–121.

Ross, S. W., Horner, R. W., & Higbee, T. (2009). Bully prevention in positive behavior support. *Journal of Applied Behavior Analysis*, *42*, 747–759.

Rouse, C. A., Everhart-Sherwood, J. M., & Alber-Morgan, S. R. (2014). Effects of self-monitoring and recruiting teacher attention on pre-vocational skills. *Education and Training in Autism and Developmental Disabilities*, *49*, 313–327.

Roxburgh, C. A., & Carbone, V. J. (2013). The effect of varying teacher presentation rates on responding during discrete trial training for two children with autism. *Behavior Modification*, *37*, 298–323.

Rozenblat, E., Brown, J. L., Brown, A. K., Reeve, S. A., & Reeve, K. F. (2009). Effects of adjusting DRO schedules on the reduction of stereotypic vocalizations in children with autism. *Behavioral Interventions*, *24*, 1–15.

Rozensky, R. H. (1974). The effect of timing of self-monitoring behavior on reducing cigarette consumption. *Journal of Consulting and Clinical Psychology*, *5*, 301–307.

Rusch, F. R., & Kazdin, A. E. (1981). Toward a methodology of withdrawal designs for the assessment of response maintenance. *Journal of Applied Behavior Analysis*, *14*, 131–140.

Ruth, W. J. (1996). Goal setting and behavioral contracting for students with emotional and behavioral difficulties: Analysis of daily, weekly, and total goal attainment. *Psychology in the Schools*, *33*, 153–158.

Ryan, C. S., & Hemmes, N. S. (2005). Effects of the contingency for homework submission on homework submission and quiz performance in a college course. *Journal of Applied Behavior Analysis*, *38*, 79–88.

Ryan, S., Ormond, T., Imwold, C., & Rotunda, R. J. (2002). The effects of a public address system on the off-task behavior of elementary physical education students. *Journal of Applied Behavior Analysis, 35*, 305–308.

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and US preschoolers. *Psychological Science, 17*, 74–81.

Saddler, P. M., & Good, E. (2006). The impact of self- and peer-grading on student learning. *Educational Assessment, 11*, 1–31.

Sagan, C. (1996). *The demon-haunted world: Science as a candle in the dark.* New York: Ballantine.

Saigh, P. A., & Umar, A. M. (1983). The effects of a good behavior game on the disruptive behavior of Sudanese elementary school students. *Journal of Applied Behavior Analysis, 16*, 339–344.

Sainato, D. M., Strain, P. S., Lefebvre, D., & Rapp, N. (1990). Effects of self-evaluation on the independent work skills of preschool children with disabilities. *Exceptional Children, 56*, 540–549.

Saini, V., Gregory, M. K., Uran, K. J., & Fantetti, M. A. (2015). Parametric analysis of response interruption and redirection as treatment for stereotypy. *Journal of Applied Behavior Analysis, 48*, 96–106.

Saini, V., Miller, S. A., & Fisher, W. W. (2016). Multiple schedules in practical application: Research trends and implications for future investigation. *Journal of Applied Behavior Analysis, 49*, 421–444.

Sajwaj, T., Culver, P., Hall, C., & Lehr, L. (1972). Three simple punishment techniques for the control of classroom disruptions. In G. Semb (Ed.), *Behavior analysis and education*. Lawrence, KS: University of Kansas.

Sakagami, T., & Lattal, K. A. (2016). The other shoe: An early operant conditioning chamber for pigeons. *The Behavior Analyst, 39*(1), 25–39.

Saksida, L. M., Raymond, S. M., & Touretzky, D. S. (1997). Shaping robot behavior using principles from instrumental conditioning. *Robotics and Autonomous Systems, 22*, 231–249.

Salend, S. J., Ellis, L. L., & Reynolds, C. J. (1989). Using self-instruction to teach vocational skills to individuals who are severely retarded. *Education and Training of the Mentally Retarded, 24*, 248–254.

Salvia, J., Ysseldyke, J. E., & Bolt, S. (2013). *Assessment in special and inclusive education* (12th ed.). Belmont, CA: Wadsworth.

Salvy, S.-J., Mulick, J. A., Butter, E., Bartlett, R. K., & Linscheid, T. R. (2004). Contingent electric shock (SIBIS) and a conditioned punisher eliminate severe head banging in preschool child. *Behavioral Interventions, 19*, 59–72.

Sanetti, L. M. H., Luiselli, J. K., & Handler, M. W. (2007). Effects of verbal and graphic performance feedback on behavior support plan implementation in a public elementary school. *Behavior Modification, 31*, 454–465. doi:10.1177/0145445506297583

Santogrossi, D. A., O'Leary, K. D., Romanczyk, R. G., & Kaufman, K. F. (1973). Self-evaluation by adolescents in a psychiatric hospital school token program. *Journal of Applied Behavior Analysis, 6*, 277–287.

Santos, P. M., Ma, M. L., & Miguel, C. F. (2015). Training intraverbal naming to establish matching-to-sample performances. *The Analysis of Verbal-Behavior, 31*, 162–182. doi.org/10.1007/s40616-015-0040-4

Sarokoff, R. A., & Strumey, P. (2004). The effects of behavioral skills training on staff implementation of discrete-trial teaching. *Journal of Applied Behavior Analysis, 37*, 535–538.

Sarokoff, R. A., Taylor, B. A., & Poulson, C. L. (2001). Teaching children with autism to engage in conversational exchanges: Script fading with embedded textual stimuli. *Journal of Applied Behavior Analysis, 34*, 81–84.

Sasso, G. M., Reimers, T. M., Cooper, L. J., Wacker, D., Berg, W., Steege, M., . . . Allaire, A. (1992). Use of descriptive and experimental analyses to identify the functional properties of aberrant behavior in school settings. *Journal of Applied Behavior Analysis, 25*, 809–821. doi:10.1901/jaba.1992.25-809

Saudargas, R. A., & Bunn, R. D. (1989). A hand-held computer system for classroom observation. *Journal of Special Education, 9*, 200–206.

Saudargas, R. A., & Zanolli, K. (1990). Momentary time sampling as an estimate of percentage time: A field validation. *Journal of Applied Behavior Analysis, 23*, 533–537.

Saunders, K. J., & Spradlin, J. E. (1989). Conditional discrimination in mentally retarded adults: The effect of training the component simple discriminations. *Journal of the Experimental Analysis of Behavior, 52*, 1–12.

Saunders, K. J., & Spradlin, J. E. (1993). Conditional discrimination in mentally retarded subjects: Programming acquisition and learning set. *Journal of the Experimental Analysis of Behavior, 60*, 571–585.

Saunders, K. J., & Williams, D. C. (1998). Stimulus-control procedures. In K. A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 193–228). New York: Plenum.

Saunders, K. J., Pilgrim, C. A., & Pennypacker, H. S. (1986). Improved proficiency of search in breast self-examination. *Cancer, 58*, 2531–2537.

Saunders, K. J., Saunders, R. R., Williams, D. C., & Spradlin, J. E. (1993). An interaction of instructions and training design on stimulus class formation: Extending the analysis of equivalence. *The Psychological Record, 43*, 725–744.

Saunders, M. D., Saunders, J. L., & Saunders, R. R. (1994). Data collection with bar code technology. In T. Thompson & D. B. Gray (Eds.), *Destructive behavior in developmental disabilities: Diagnosis and treatment* (pp. 102–116). Thousand Oaks, CA: Sage.

Saunders, R. R., & Green, G. (1999). A discrimination analysis of training-structure effects on stimulus equivalence outcomes. *Journal of the Experimental Analysis of Behavior, 72*, 117–137.

Saunders, R. R., & Sailer, W. (1979). A comparison of three strategies of reinforcement on two-choice learning problems with severely retarded children. *AAESPH Review, 4*, 323–333.

Saunders, R. R., Chaney, L., & Marquis, J. G. (2005). Equivalence class establishment with two-, three-, and four-choice matching to sample by senior citizens. *The Psychological Record, 55*, 539–559.

Saunders, R. R., Saunders, K. J., Kirby, K. C., & Spradlin, J. E. (1988). The merger of equivalence classes by unreinforced conditional selection of comparison stimuli. *Journal of the Experimental Analysis of Behavior, 50*, 145–162.

Saunders, R. R., Wachter, J. A., & Spradlin, J. E. (1988). Establishing auditory stimulus control over an eight-member equivalence class via conditional discrimination procedure. *Journal of the Experimental Analysis of Behavior, 49*, 95–115.

Sautter, R. A., & LeBlanc, L. A. (2006). Empirical applications of Skinner's analysis of verbal behavior with humans. *The Analysis of Verbal Behavior, 22*, 35–48.

Sautter, R. A., LeBlanc, L. A., Jay, A. A., Goldsmith, T. R., & Carr, J. E. (2011). The role of problem solving in complex intraverbal repertoires. *Journal of Applied Behavior Analysis, 44*, 227–244. doi.org/10.1901/jaba.2011.44-227

Savage, M. N. (2014). Self-operated auditory prompting systems: Creating and using them to support students with disabilities. *Teaching Exceptional Children, 47*(1), 46–55.

Savage, T. (1998). Shaping: The link between rats and robots. *Connection Science, 10*(3/4), 321–340.

Sawyer, M. R., Andzik, N. R., Kranak, M. P., Willke, C. P., Curiel, E. S. L., Hensley, L. E., & Neef, N. A. (2017). Improving pre-service teachers' performance skills through behavioral skills training. *Behavior Analysis in Practice, 10*, 296–300.

Scheithauer, M., O'Connor, J., & Toby, L. M. (2015). Assessment of self-restraint using a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 48*, 907–911.

Scheithauer, M. C., & Tiger, J. H. (2014). Teaching braille line tracking using stimulus fading. *Journal of Applied Behavior Analysis, 47*, 612–616.

Schepis, M. M., Reid, D. H., Behrmann, M. M., & Sutton, K. A. (1998). Increasing communicative interactions of young children with autism using a voice output communication aid and naturalistic teaching. *Journal of Applied Behavior Analysis, 31*, 561–578.

Scherrer, M. D., & Wilder, D. A. (2008). Training to increase safe tray carrying among cocktail servers. *Journal of Applied Behavior Analysis, 41*, 131–135. doi:10.1901/jaba.2008.41-131

Scheuermann, B., & Webber, J. (1996). Level systems: Problems and solutions. *Beyond Behavior, 7*, 12–17.

Schleien, S. J., Wehman, P., & Kiernan, J. (1981). Teaching leisure skills to severely handicapped adults: An age-appropriate darts game. *Journal of Applied Behavior Analysis, 14*, 513–519.

Schlichenmeyer, K. J., Dube, W. V., & Vargas-Irwin, M. (2015). Stimulus fading and response elaboration in differential reinforcement for alternative behavior. *Behavioral Interventions, 30*, 51–64. doi:10.1002/bin.1402

Schlinger, H. D. (2007). Does love or differential reinforcement open doors? *Journal of Applied Behavior Analysis*, *40*, 385. doi:10.1901/jaba.2007.209-06

Schlinger, H. D. (2008a). Consciousness is nothing but a word. *Skeptic*, *4*, 58–63.

Schlinger, H. D. (2008b). Listening is behaving verbally. *The Behavior Analyst*, *31*, 145–161.

Schlinger, H. D. (2011). Introduction: Private events in a natural science of behavior. *The Behavior Analyst*, *34*, 181–184.

Schlinger, H. D. (2011). Skinner as missionary and prophet: A review of *Burrhus F. Skinner: Shaper of Behaviour*. *Journal of Applied Behavior Analysis*, *44*, 217–225.

Schlinger, H. D. (2017). The importance of analysis in applied behavior analysis. *Behavior Analysis: Research and Practice*, *17*, 334–346. http://dx.doi.org/10.1037/bar0000080

Schlinger, H. D. (2018, Issue 1). A letter to the *New York Times*. *Operants*, 10–11.

Schlinger, H. D., & Normand, M. P. (2013). On the origin and functions of the term *functional analysis*. *Journal of Applied Behavior Analysis*, *46*, 285–288.

Schlinger, H. D., Derenne, A., & Baron, A. (2008). What 50 years of research tell us about pausing under ratio schedules of reinforcement. *The Behavior Analyst*, *31*, 39–60.

Schlinger, H. D., Jr. (1995). *A behavior analytic view of child development*. New York: Plenum Press.

Schlinger, H. D., Jr. (2008). Conditioning the behavior of the listener. *International Journal of Psychology and Psychological Therapy*, *8*, 309–322.

Schlinger, H. D., Jr., & Blakely, E. (1987). Function-altering effects of contingency-specifying stimuli. *The Behavior Analyst*, *10*, 41–45.

Schlinger, H. D., Jr., & Blakely, E. (1994). A descriptive taxonomy of environmental operations and its implications for behavior analysis. *The Behavior Analyst*, *17*, 43–57.

Schmidt, A. C., Hanley, G. P., & Layer, S. A. (2009). A further analysis of the value of choice: Controlling for illusory discriminative stimuli and evaluating the effects of less preferred items. *Journal of Applied Behavior Analysis*, *42*, 711–716

Schneider, S. M. (2012). *The science of consequences: How they affect genes, change the brains, and impact our world*. Amherst, NY: Prometheus.

Schnelle, J. F., Kirchner, R. E., Macrae, J. W., McNees, M. P., Eck, R. H., Snodgrass, S., . . . Uselton, P. H., Jr. (1979). Program evaluation research: An experimental cost-effectiveness analysis of an armed robbery intervention program. *Journal of Applied Behavior Analysis*, *12*, 615–623.

Schoenfeld, W. N. (1995). "Reinforcement" in behavior theory. *The Behavior Analyst*, *18*, 173–185.

Schramm, R. (2011). *Motivation and reinforcement: Turning the tables on autism*. Hespe, Germany: Knospe-ABA.

Schreck, K. A., & Mazur, A. (2008). Behavior analyst use of and beliefs in treatments for people with autism. *Behavior Interventions*, 23, 201–212.

Schreck, K. A., & Miller, V. A. (2010). How to behave ethically in a world of fads. *Behavior Interventions*, *25*, 307–324.

Schultz, N. R., Kohn, C. S., & Musto, A. (2017). Examination of a multi-element intervention on college students' electricity consumption in on-campus housing. *Behavioral Interventions*, *32*, 79–90. doi:10.1002/bin.1463

Schumm, J. S., Vaughn, D., Haager, D., McDowell, J., Rothlein, L., & Saumell, L. (1995). General education teacher planning: What can students with learning disabilities expect? *Exceptional Children*, *61*, 335–352.

Schuster, J. W., Griffen, A. K., & Wolery, M. (1992). Comparison of simultaneous prompting and constant time delay procedures in teaching sight words to elementary students with moderate mental retardation. *Journal of Behavioral Education*, *7*, 305–325.

Schwanhausser, L. (2009). Application of the mindfulness-acceptance-commitment (MAC) protocol with an adolescent springboard diver. *Journal of Clinical Sport Psychology*, *3*, 377–395.

Schwartz, B. (1974). On going back to nature: A review of Seligman and Hager's *Biological Boundaries of Learning*. *Journal of the Experimental Analysis of Behavior*, *21*, 183–198.

Schwartz, B. (1989). *Psychology of learning and behavior*. New York: Norton.

Schwarz, M. L., & Hawkins, R. P. (1970). Application of delayed reinforcement procedures to the behavior of an elementary school child. *Journal of Applied Behavior Analysis*, *3*, 85–96.

Schweitzer, J. B., & Sulzer-Azaroff, B. (1988). Self-control: Teaching tolerance for delay in impulsive children. *Journal of the Experimental Analysis of Behavior*, *50*, 173–186. doi:10.1901/jeab.1988.50-173

Scott, D., Scott, L. M., & Goldwater, B. (1997). A performance improvement program for an international-level track and field athlete. *Journal of Applied Behavior Analysis*, *30*, 573–575.

Seaver, J. L., & Bourret, J. C. (2014). An evaluation of response prompts for teaching behavior chains. *Journal of Applied Behavior Analysis*, *47*, 777–792.

Sella, A. C., Ribeiro, D. M., & White, G. W. (2014). Effects of an online stimulus equivalence teaching procedure on research design open-ended questions performance of international undergraduate students. *The Psychological Record*, *64*, 89–103.

Sellers, T. P., Alai-Rosales, S., & MacDonald, R. P. F. (2016). Taking full responsibility: The ethics of supervision in behavior analytic practice. *Behavior Analysis in Practice*, *9*, 299–308. doi 10.1007/s40617-016-0144-x

Sellers, T. P., LeBlanc, L. A., & Valentino, A. L. (2016). Recommendations for detecting and addressing barriers to successful supervision. *Behavior Analysis in Practice*, *9*, 309–319. doi 10.1007/s40617-016-0142-z

Sellers, T. P., Valentino, A. L., & LeBlanc, L. A. (2016). Recommended practices for individual supervision of aspiring behavior analysts. *Behavior Analysis in Practice*, *9*, 274–286. doi 10.1007/s40617-016-0110-7

Seraganian, P., & vom Saal, W. (1969). Blocking the development of stimulus control when stimuli indicate periods of nonreinforcement. *Journal of the Experimental Analysis of Behavior*, *12*, 767–772.

Seymour, F. W., & Stokes, T. F. (1976). Self-recording in training girls to increase work rate and evoke staff praise in an institution for offenders. *Journal of Applied Behavior Analysis*, *9*, 41–54.

Seymour, M. A. (2002). Case study: A retired athlete runs down comeback road. In R. W. Malott & H. Harrison (Eds.), *I'll stop procrastinating when I get around to it: Plus other cool ways to succeed in school and life using behavior analysis to get your act together* (pp. 7–12). Kalamazoo, MI: Department of Psychology, Western Michigan University.

Shaddish, W. R. (2014). Analysis and meta-analysis of single-case designs: An introduction. *Journal of School Psychology*, *52*, 109-122. doi.org/10.1016/j.jsp.2013.11.009

Shaddish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference* (2nd ed.). Boston: Cengage Learning.

Shafer, E. (1993). Teaching topography-based and stimulus selection-based verbal behavior to developmentally disabled individuals: Some considerations. *The Analysis of Verbal Behavior*, *11*, 117–133.

Shahan, T. A. (2010). Conditioned reinforcement and response strength. *Journal of the Experimental Analysis of Behavior*, *93*, 269–289.

Shahan, T. A. (2013). Attention and conditioned reinforcement. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. 387–410). Washington, DC: American Psychological Association.

Shahan, T. A., & Chase, P. N. (2002). Novelty, stimulus control, and operant variability. *The Behavior Analyst*, *25*, 175–190.

Shahan, T. A., & Sweeney, M. M. (2011). A model of resurgence based on behavioral momentum theory. *Journal of the Experimental Analysis of Behavior*, *95*, 91–108.

Shapiro, M., Kazemi, E., Pogosjana, M., Rios, D., & Mendoza, M. (2016). Preference assessment training via self-instruction: A replication and extension. *Journal of Applied Behavior Analysis*, *49*, 794–808.

Sharp, W. G., Odom, A., & Jaquess, D. L. (2012). Comparison of upright and flipped spoon presentations to guide treatment of food refusal. *Journal of Applied Behavior Analysis*, *45*, 83–96. doi.org/10.1901/jaba.2012.45-83

Shawler, L. A., & Miguel, C. (2015). The effects of motor and vocal response interruption and redirection on vocal stereotypy and appropriate vocalizations. *Behavioral Interventions*, *30*, 112–134. doi:10.1002/bin.1407

Sherer, M., Pierce, K. L., Paredes, S., Kisacky, K. L., Ingersoll, B., & Schreibman, L. (2001). Enhancing conversation skills in children with autism via video technology: Which is better, "Self" or "Other" as a model? *Behavior Modification*, *25*, 140–158.

Shermer, M. (1997). *Why people believe weird things*. New York: Freeman.

Shermer, M. (2002). *Why people believe weird things: Pseudoscience, superstition, and other confusions of our time*. New York: Henry Holt.

Shillingsburg, M. A., & Valentino, A. L. (2011). Teaching a child with autism to mand for information using "how." *The Analysis of Verbal Behavior*, *27*, 179–184.

Shillingsburg, M. A., Bowen, C. N., Valentino, A. L., & Pierce, L. E. (2014). Mands for information using "who?" and "which?" in the presence of

establishing and abolishing operations. *Journal of Applied Behavior Analysis, 47*, 136–150.

Shillingsburg, M. A., Gayman, C. M., & Walton, W. (2016). Using textual prompts to teach mands for information using "Who?" *The Analysis of Verbal Behavior, 32*, 1–14. doi.org/10.1007/s40616-016-0053-7

Shillingsburg, M. A., Hollander, D. L., Yosick, R. N., Bowen, C., & Muskat, L. R. (2015). Stimulus-stimulus pairing to increase vocalizations in children with language delays: A review. *The Analysis of Verbal Behavior, 31*, 215–235.

Shillingsburg, M. A., Valentino, A. L., Bowen, C. N., Bradley, D., & Zavatkay, D. (2011). Teaching children with autism to request information. *Research in Autism Spectrum Disorders, 5*, 670–679.

Shimabukuro, S. M., Pratter, M. A., Jenkins, A., & Edelen-Smith, P. (2000). The effects of self-monitoring of academic performance on students with learning disabilities and ADD/ADHD. *Education and Treatment of Children, 22*, 397–414.

Shimmel, S. (1977). Anger and its control in Greco-Roman and modern psychology. *Psychiatry, 42*, 320–327.

Shimmel, S. (1979). Free will, guilt, and self-control in rabbinic Judaism and contemporary psychology. *Judaism, 26*, 418–429.

Shimoff, E., & Catania, A. C. (1995). Using computers to teach behavior analysis. *The Behavior Analyst, 18*, 307–316.

Shinn, M. M., & Shinn, N. R. (2002). *AIMSweb® training workbook: Administration and scoring of reading maze for use in general outcome measurement*. Eden Prairie, MN: NCS Pearson.

Shipley-Benamou, R., Lutzker, J. R., & Taubman, M. (2002). Teaching daily living skills to children with autism through instructional video modeling. *Journal of Positive Behavior Interventions, 4*, 166–177.

Shirley, M. J., Iwata, B. A., Kahng, S. W., Mazaleski, J. L., & Lerman, D. C. (1997). Does functional communication training compete with ongoing contingencies of reinforcement? An analysis during response acquisition and maintenance. *Journal of Applied Behavior Analysis, 30*, 93–104.

Shook, G. L., & Favell, J. E. (1996). Identifying qualified professionals in behavior analysis. In Maurice, C., Green, G., & Luce, S. C. (Eds.), *Behavioral intervention for young children with autism: A manual for parents and professionals* (pp. 221–229). Austin, TX: PRO-ED.

Shook, G. L., & Neisworth, J. (2005). Ensuring appropriate qualifications for applied behavior analysis professionals: The Behavior Analyst Certification Board. *Exceptionality, 13*(1), 3–10.

Shook, G. L., Johnston, J. M., & Mellichamp, F. (2004). Determining essential content for applied behavior analyst practitioners. *The Behavior Analyst, 27*, 67–94.

Shore, B. A., Iwata, B. A., DeLeon, I. G., Kahng, S., & Smith, R. G. (1997). An analysis of reinforcer substitutability using object manipulation and self-injury as competing responses. *Journal of Applied Behavior Analysis, 30*, 439–449.

Shrewsberry, R. D. (1977). *Assignment completion in group parent training* (Doctoral dissertation). Retrieved from ProQuest Dissertations Publishing, 7731977.

Shriberg, J. D. (1971). A system for monitoring and conditioning modal fundamental frequency of speech. *Journal of Applied Behavior Analysis, 4*, 565–568.

Sidenar, T. M., Shabani, D. B., & Carr, J. E. (2004). A review of the behavioral evaluation strategy taxonomy (Best) software application. *Behavioral Interventions, 19*, 275–285.

Sidener, D. W. (2006). Joint control for dummies*: An elaboration of Lowenkron's model of joint (stimulus) control. *The Analysis of Verbal Behavior, 22*(1), 119–122.

Sidener, D. W., & Michael, J. (2006). Generalization of relational matching to sample in children: A direct replication. *The Analysis of Verbal Behavior, 22*, 171–181. 10.1007/BF03393037

Sidener, T. M., Shabani, D. B., Carr, J. E., & Roland, J. P. (2006). An evaluation of strategies to maintain mands at practical levels. *Research in Developmental Disabilities, 27*, 632–644.

Sidman, M. (1953). Two temporal parameters in the maintenance of avoidance behavior by the white rat. *Journal of Comparative and Physiological Psychology, 46*, 253–261.

Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. New York: Basic Books/Boston: Authors Cooperative (reprinted).

Sidman, M. (1960). *Tactics of scientific research*. New York: Basic Books.

Sidman, M. (1962). Classic avoidance without a warning stimulus. *Journal of the Experimental Analysis of Behavior, 5*, 97–104.

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research, 14*, 5–13.

Sidman, M. (1980). A note on the measurement of conditional discrimination. *Journal of the Experimental Analysis of Behavior, 33*, 285–289.

Sidman, M. (1981). Remarks. *Behaviorism, 9*, 127–129.

Sidman, M. (1986a). The measurement of behavioral development. In N. Krasnegor, D. B. Gray, & T. Thompson (Eds.), *Advances in behavior pharmacology: Vol. 5: Developmental behavioral pharmacology* (pp. 43–52). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sidman, M. (1986b). Functional analysis of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 213–245). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sidman, M. (1987). Two choices are not enough. *Behavior Analysis, 22*, 11–18.

Sidman, M. (1992). Equivalence relations: Some basic considerations. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 15–27). Reno, NV: Context Press.

Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston: Authors Cooperative.

Sidman, M. (2000). Applied behavior analysis: Back to basics. *Behaviorology, 5* (1), 15–37.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74*, 127–146.

Sidman, M. (2002). Notes from the beginning of time. *The Behavior Analyst, 25*, 3–13.

Sidman, M. (2006). The distinction between positive and negative reinforcement: Some additional considerations. *The Behavior Analyst, 29*, 135–139.

Sidman, M. (2007). The analysis of behavior: What's in it for us? *Journal of the Experimental Analysis of Behavior, 87*, 309–316.

Sidman, M. (2009). Equivalence relations and behavior: An introductory tutorial. *The Analysis of Verbal Behavior, 25*, 5–17.

Sidman, M. (2013). Foreword. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. xv–xvii). Washington, DC: American Psychological Association.

Sidman, M., & Cresson, O. (1973). Reading and crossmodal transfer of stimulus equivalences in severe retardation. *American Journal of Mental Deficiency, 77*, 515–523.

Sidman, M., & Stoddard, L. T. (1967). The effectiveness of fading in programming a simultaneous form discrimination for retarded children. *Journal of the Experimental Analysis of Behavior, 10*, 3–15.

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37*, 5–22. doi.org/10.1901/jeab.1982.37-5

Sidman, M., Cresson, O., Jr., & Willson-Morris, M. (1974). Acquisition of matching-to-sample via mediated transfer. *Journal of the Experimental Analysis of Behavior, 22*, 261–273.

Sidman, M., Kirk, B., & Willson-Morris, M. (1985). Six-member stimulus classes generated by conditional-discrimination procedures. *Journal of the Experimental Analysis of Behavior, 43*, 21–42.

Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children. *Journal of the Experimental Analysis of Behavior, 37*, 23–44.

Sidman, M., Willson-Morris, M., & Kirk, B. (1986). Matching-to-sample procedures and the development of equivalence relations: The role of naming. *Analysis and Intervention in Developmental Disabilities, 6*, 1–19.

Sidman, M., Wynne, C. K., Maquire, R. W., & Barnes, T. (1989). Functional classes and equivalence relations. *Journal of the Experimental Analysis of Behavior, 52*, 261–274.

Sigafoos, J., & Saggers, E. (1995). A discrete-trial approach to the functional analysis of aggressive behaviour in two boys with autism. *Australia and New Zealand Journal of Developmental Disabilities, 4*, 287–297.

Sigafoos, J., O'Reilly, M., Cannella, H., Upadhyaya, M., Edrisinha, C., Lancioni, G. E., . . . Young, D. (2005). Computer-presented video prompting for teaching microwave oven use to three adults with developmental disabilities. *Journal of Behavioral Education, 14*(3), 189–201.

Sigurdsson, S. O., & Austin, J. (2008). Using real-time visual feedback to improve posture at computer workstations. *Journal of Applied Behavior Analysis, 41*, 365–375. doi:10.1901/jaba.2008.41-365

Silverman, K., Kaminski, B. J., Higgins, S. T., & Brady, J. V. (2011). Behavior analysis and treatment of drug addiction. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 451–471). New York: Guilford Press.

Silvestri, S. M. (2004). *The effects of self-scoring on teachers' positive statements during classroom instruction*. Unpublished doctoral dissertation, The Ohio State University, Columbus, OH.

Silvestri, S. M. (2005). *How to make a graph using Microsoft Excel*. Unpublished manuscript. Columbus, OH: The Ohio State University.

Silvestri, S. M., & Heward, W. L. (2016). The neutralization of special education, revisited. In R. M. Foxx & J. A. Mulick (Eds.), *Controversial therapies for autism and intellectual disabilities: Fads, fashion, and science in professional practice* (2nd ed.) (pp. 136–153). New York: Routledge.

Simek, T. C., O'Brien, R. M., & Figlerski, L. B. (1994). Contracting and chaining to improve the performance of a college golf team: Improvement and deterioration. *Perceptual and Motor Skills, 78*(3), 1099.

Simonsen, B., MacSuga, A. S., Fallon, L. M., & Sugai, G. (2013). The effects of self-monitoring on teachers' use of specific praise. *Journal of Positive Behavior Interventions, 15*, 5–15.

Simpson, M. J. A., & Simpson, A. E. (1977). One-zero and scan method for sampling behavior. *Animal Behavior, 25*, 726–731.

Sindelar, P. T., Rosenberg, M. S., & Wilson, R. J. (1985). An adapted alternating treatments design for instructional research. *Education and Treatment of Children, 8*, 67–76.

Singer, G., Singer, J., & Horner, R. (1987). Using pretask requests to increase the probability of compliance for students with severe disabilities. *Journal of the Association for Persons with Severe Handicaps, 12*, 287–291.

Singh, N. N., & Singh, J. (1986). Increasing oral reading proficiency: A comparative analysis of drill and positive practice overcorrection procedures. *Behavior Modification, 10*, 115–130.

Singh, N. N., & Winton, A. S. (1985). Controlling pica by components of an overcorrection procedure. *American Journal of Mental Deficiency, 90*, 40–45.

Singh, N. N., Dawson, M. J., & Manning, P. (1981). Effects of spaced responding DRI on the stereotyped behavior of profoundly retarded persons. *Journal of Applied Behavior Analysis, 14*, 521–526.

Singh, N. N., Lancioni, G. E., Manikam, R., Winton, A. S., Singh, A. N., Singh, J., & Singh, A. D. (2011). A mindfulness-based strategy for self-management of aggressive behavior in adolescents with autism. *Research in Autism Spectrum Disorders, 5*, 1153–1158.

Singh, N. N., Lancioni, G. E., Winton, A. S., Karazsia, B. T., Myers, R. E., Latham, L. L., & Singh, J. (2014). Mindfulness-based positive behavior support (MBPBS) for mothers of adolescents with autism spectrum disorder: Effects on adolescents' behavior and parental stress. *Mindfulness, 5*, 646–657.

Singh, N. N., Singh, J., & Winton, A. S. (1984). Positive practice overcorrection of oral reading errors. *Behavior Modification, 8*, 23–37.

Sisson, L. A., & Barrett, R. P. (1984). Alternating treatments comparison of oral and total communication training with minimally verbal retarded children. *Journal of Applied Behavior Analysis, 17*, 559–566.

Sisson, L. A., Van Hasselt, V. B., Hersen, M., & Aurand, J. C. (1988). Tripartite behavioral intervention to reduce stereotypic and disruptive behaviors in young multi-handicapped children. *Behavior Therapy, 19*, 503–526.

Sivaraman, M. (2017). Using multiple exemplar training to teach empathy skills to children with autism. *Behavior Analysis in Practice, 10*, 337–346. doi:10.1007/s40617-017-0183-y

Skiba, R., & Raison, J. (1990). Relationship between the use of timeout and academic achievement. *Exceptional Children, 57*(1), 36–46.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1945). The operational analysis of psychological terms. *Psychological Review, 52*, 270–277.

Skinner, B. F. (1948). Superstition in the pigeon. *Journal of Experimental Psychology, 38*, 168–172.

Skinner, B. F. (1948). *Walden two*. New York: Macmillan.

Skinner, B. F. (1953). *Science and human behavior*. New York: Free Press.

Skinner, B. F. (1956). A case history in scientific method. *American Psychologist, 11*, 221–233.

Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1958). Teaching machines. *Science, 128*, 969–977.

Skinner, B. F. (1966). Operant behavior. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 12–32). New York: Appleton-Century-Crofts.

Skinner, B. F. (1967). B. F. Skinner: An autobiography. In E. G. Boring & G. Lindzey (Eds.), *A history of psychology in autobiography* (Vol. 5, pp. 387–413). New York: Irvington.

Skinner, B. F. (1968). *The technology of teaching*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1969). *Contingencies of reinforcement: A theoretical analysis*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1971). *Beyond freedom and dignity*. New York: Knopf.

Skinner, B. F. (1974). *About behaviorism*. New York: Knopf.

Skinner, B. F. (1976). *Particulars of my life*. Washington Square, NY: New York University Press.

Skinner, B. F. (1978). *Reflections on behaviorism and society*. Englewood Cliffs, NJ: Prentice Hall.

Skinner, B. F. (1978). The shaping of phylogenic behavior. In B. F. Skinner, *Reflections on behaviorism and society* (pp. 163–170). Englewood Cliffs, NJ: Prentice Hall. (Original work published 1975).

Skinner, B. F. (1979). *The shaping of a behaviorist*. Washington Square, NY: New York University Press.

Skinner, B. F. (1981a). Selection by consequences. *Science, 213*, 501–504.

Skinner, B. F. (1981b). How to discover what you have to say—A talk to students. *The Behavior Analyst, 4*, 1–7.

Skinner, B. F. (1982). Contrived reinforcement. *The Behavior Analyst, 5*, 3–8.

Skinner, B. F. (1983). *A matter of consequences: Part three of an autobiography*. Washington Square, NY: New York University Press.

Skinner, B. F. (1989). *Recent issues in the analysis of behavior*. Columbus, OH: Merrill.

Skinner, C. H., Cashwell, T. H., & Skinner, A. L. (2000). Increasing tootling: The effects of a peer-mediated group contingency program on students' reports of peers' prosocial behavior. *Psychology in the Schools, 37*(3), 263–270.

Skinner, C. H., Fletcher, P. A., Wildmon, M., & Belfiore, P. J. (1996). Improving assignment preference through interspersing additional problems: Brief versus easy problems. *Journal of Behavioral Education, 6*, 427–436.

Skinner, C. H., Skinner, A. L., & Burton, B. (2009). Applying group-oriented contingencies in the classroom. In A. Akin-Little, S. G. Little, M. A. Bray, & T. J. Kehle (Eds.), *Behavioral interventions in schools: Evidence-based positive strategies* (pp. 157–170). Washington, DC: APA Press.

Skinner, C. H., Skinner, C. F., Skinner, A. L., & Cashwell, T. H. (1999). Using interdependent contingencies with groups of students: Why the principal kissed a pig. *Educational Administration Quarterly, 35* (supplemental), 806–820. https://doi.org/10.1177/00131619921968833

Slaton, J. D., & Hanley, G. P. (2016). Effects of multiple versus chained schedules on stereotypy and item engagement. *Journal of Applied Behavior Analysis, 49*, 927–946. doi:10.1002/jaba.345

Slattery, B., & Stewart, I. (2014). Hierarchical classification as relational framing. *Journal of the Experimental Analysis of Behavior, 101*, 61–75.

Slattery, B., Stewart, I., & O'Hora, D. (2011). Testing for transitive class containment as a feature of hierarchical classification. *Journal of the Experimental Analysis of Behavior, 96*, 243–260.

Sleeper, J. D., LeBlanc, L. A., Mueller, J., Valentino, A. L., Fazzio, D., & Raetz, P. B. (2017). The effects of electronic data collection on the percentage of current clinician graphs and organizational return on investment. *Journal of Organizational Behavior Management*, doi:10.1080/01608061.2016.1267065.

Slocum, S. K., & Tiger, J. H. (2011). An assessment of the efficiency of and child preference for forward and backward chaining. *Journal of Applied Behavior Analysis, 44*, 793–805.

Slocum, S. K., & Vollmer, T. R. (2015). A comparison of positive and negative reinforcement for compliance to treat problem behavior maintained by escape. *Journal of Applied Behavior Analysis, 48*, 563–574.

Slocum, S. K., Mehrkam, L. R., Peters, K. P., & Vollmer, T. R. (2017). Using differential reinforcement of a discard response to treat pica. *Behavioral Interventions, 32*, 234–241. doi:10.1002/bin.1483

Sloman, K. N., Vollmer, T. R., Cotnoir, N. M., Borrero, C. S. W., Borrero, J. C., Samaha, A., & St. Peter, C. C. (2005). Descriptive analysis of caregiver reprimands. *Journal of Applied Behavior Analysis, 38*, 373–383.

Smeets, P. M., & Barnes-Holmes, D. (2003). Children's emergent preferences for soft drinks: Stimulus-equivalence and transfer. *Journal of Economic Psychology, 24*, 603–618.

Smith, D. H. (1987). Telling stories as a way of doing ethics. *Journal of the Florida Medical Association, 74*, 581–588.

Smith, D. H. (1993). Stories, values, and patient care decisions. In C. Conrad (Ed.), *Ethical nexus* (pp. 123–148). Norwood, NJ: Ablex.

Smith, D. P., Eikeseth, S., Fletcher, S. E., Montebelli, L., Smith, H. R., & Taylor, J. C. (2016). Emergent intraverbal forms may occur as a result of listener training for children with autism. *The Analysis of Verbal Behavior, 32*, 27–37. doi.org/10.1007/s40616-016-0057-3

Smith, L. D. (1992). On prediction and control: B. F. Skinner and the technological ideal of science. *American Psychologist, 47*, 216–223.

Smith, L. D., Best, L. A., Stubbs, D. A., Johnston, J., & Bastiani-Archibald, A. (2000). Scientific graphs and the hierarchy of the sciences: A Latourian survey of inscription practices. *Social Studies of Science, 30*, 73–94.

Smith, N. G., & Morris, E. K. (2018). *An expanded bibliography of the published primary-source works of B. F. Skinner: With references to works reprinted in Skinner's collections and texts*. Cambridge, MA: B. F. Skinner Foundation. Retrieved August 11, 2018, from https://www.bfskinner.org/publications/full-bibliography/

Smith, R. G. (2011). Developing antecedent interventions for problem behavior. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 297–316). New York: Guilford Press.

Smith, R. G., & Churchill, R. M. (2002). Identification of environmental determinants of behavior disorders through functional analysis of precursor behaviors. *Journal of Applied Behavior Analysis, 35*, 125–136. doi:10.1901/jaba.2002.35-125

Smith, R. G., & Iwata, B. A. (1997). Antecedent influences of behavior disorders. *Journal of Applied Behavior Analysis, 30*, 343–375.

Smith, R. G., Iwata, B. A., Goh, H., & Shore, B. A. (1995). Analysis of establishing operations for self-injury maintained by escape. *Journal of Applied Behavior Analysis, 28*, 515–535.

Smith, R. G., Lerman, D. C., & Iwata, B. A. (1996). Self-restraint as positive reinforcement for self-injurious behavior. *Journal of Applied Behavior Analysis, 29*, 99–102.

Smith, R. G., Russo, L., & Le, D. D. (1999). Distinguishing between extinction and punishment effects of response blocking: A replication. *Journal of Applied Behavior Analysis, 32*, 367–370.

Smith, S., & Farrell, D. (1993). Level system use in special education: Classroom intervention with prima facie appeal. *Behavioral Disorders, 18*(4), 251–264.

Smith, T. (2001). Discrete trial training in the treatment of autism. *Focus on Autism and Other Developmental Disabilities, 16*, 86–92. doi:10.1177/108835760101600204

Smith, T., Eikeseth, S., Klevstrand, M., & Lovaas, O. I. (1997). Intensive behavioral treatment for preschoolers with severe mental retardation and pervasive developmental disorders. *American Journal on Mental Retardation, 102*, 238–249.

Smith, T., Groen, A. D., & Wynn, J. W. (2000). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American Journal on Mental Retardation, 105*, 269–285.

Smyth, S., Barnes-Holmes, D., & Barnes-Holmes, Y. (2008). Acquired equivalence in human discrimination learning: The role of propositional knowledge. *Journal of Experimental Psychology: Animal Behavior Processes, 34*, 167–177.

Snodgrass, M. R., Chung, M. Y., Meadan, H., & Halle, J. W. (2018). Social validity in single-case research: A systematic literature review of prevalence and application. *Research in Developmental Disabilities, 74*, 160–173.

Solanto, M. V., Jacobson, M. S., Heller, L., Golden, N. H., & Hertz, S. (1994). Rate of weight gain of inpatients with anorexia nervosa under two behavioral contracts. *Pediatrics, 93*(6), 989–991.

Solnick, J. V., Rincover, A., & Peterson, C. R. (1977). Some determinants of the reinforcing and punishing effects of timeout. *Journal of Applied Behavior Analysis, 10*, 415–424.

Solomon, B. G., Howard, T. K., & Stein, B. L. (2015). Critical assumptions and distribution features pertaining to contemporary single-case effect sizes. *Journal of Behavioral Education, 24*, 438–458. doi 10.1007/s10864-015-9221-4

Solomon, R. L. (1964). Punishment. *American Psychologist, 19*, 239–253.

Solomon, R. L., Kamin, L. J., & Wynne, L. C. (1953). Traumatic avoidance learning: The outcomes of several extinction procedures with dogs. *The Journal of Abnormal and Social Psychology, 48*(2), 291–302.

Southall, C., & Gast, D. L. (2011). Self-management procedures: A comparison across the autism spectrum. *Education and Training in Autism and Developmental Disabilities, 46*, 155–171.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (2005). *Vineland Adaptive Behavior Scales: Second edition (Vineland—II)*. Upper Saddle River, NJ: Pearson Assessments.

Speckman, J., Greer, R. D., & Rivera-Valdes, C. (2012). Multiple exemplar instruction and the emergence of generative production of suffixes as autoclitic frames. *The Analysis of Verbal Behavior, 28*, 83–99. doi.org/10.1007/BF03393109

Speelman, R. C., Whiting, S. W., & Dixon, M. R. (2015). Using behavioral skill training and video rehearsal to teach blackjack skills. *Journal of Applied Analysis of Behavior, 48*, 632–642. doi:10.1002/jaba.225

Spek, A. A., Van Ham, N. C., & Nyklíček, I. (2013). Mindfulness-based therapy in adults with an autism spectrum disorder: A randomized controlled trial. *Research in developmental disabilities, 34*, 246–253.

Spieler, C., & Miltenberger, R. (2017). Using awareness training to decrease nervous habits during public speaking. *Journal of Applied Behavior Analysis, 50*, 38–47. doi:10.1002/jaba.362

Spooner, F., Spooner, D., & Ulicny, G. R. (1986). Comparisons of modified backward chaining: Backward chaining with leaps ahead and reverse chaining with leaps ahead. *Education and Treatment of Children, 9*(2), 122–134.

Spradlin, J. E. (1963). Assessment of speech and language of retarded children: The Parsons language sample. *Journal of Speech and Hearing Disorders Monograph, 10*, 8–31.

Spradlin, J. E. (1966). Environmental factors and the language development of retarded children. In S. Rosenberg (Ed.), *Developments in applied psycholinguist research* (pp. 261–290). Riverside, NJ: The Macmillan Company.

Spradlin, J. E. (2002). Punishment: A primary response. *Journal of Applied Behavior Analysis, 35*, 475–477.

Spradlin, J. E., & Simon, J. L. (2011). Stimulus control and generalization. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds). *Handbook of applied behavior analysis* (pp. 76–91). New York: Guilford Press.

Spradlin, J. E., Cotter, V. W., & Baxley, N. (1973). Establishing a conditional discrimination without direct training: A study of transfer with retarded adolescents. *American Journal of Mental Deficiency, 77*, 556–566.

Sprague, J., & Walker, H. (2000). Early identification and intervention for youth with antisocial and violent behavior. *Exceptional Children, 66*, 367–379.

Sprague, J. R., & Horner, R. H. (1984). The effects of single instance, multiple instance, and general case training on generalized vending machine used by moderately and severely handicapped students. *Journal of Applied Behavior Analysis, 17*, 273–278.

Sprague, J. R., & Horner, R. H. (1990). Easy does it: Preventing challenging behaviors. *Teaching Exceptional Children, 23*(1), 13–15.

Sprague, J. R., & Horner, R. H. (1991). Determining the acceptability of behavior support plans. In M. Wang, H. Walberg, & M. Reynolds (Eds.), *Handbook of special education* (pp. 125–142). Oxford, London, UK: Pergamon Press.

Spreat, S. (1982). Weighing treatment alternatives: Which is less restrictive? *Woodhaven Center E & R Technical Report, 82-11(1)*. Philadelphia: Temple University.

Spreat, S., & Connelly, L. (1996). Reliability analysis of the Motivation Assessment Scale. *American Journal on Mental Retardation, 100*, 528–532.

Sprinkle, E. C., & Miguel, C. F. (2012). The effects of listener and speaker training on emergent relations in children with autism. *The Analysis of Verbal Behavior, 28*, 111–117. doi.org/10.1007/BF03393111

Squires, J., Wilder, D. A., Fixsen, A., Hess, E., Rost, K., Curran, R., & Zonneveld, K. (2008). The effects of task clarification, visual prompts, and graphic feedback on customer greeting and up-selling in a restaurant. *Journal of Organizational Behavior Management, 27*, 1–13. doi.org/10.1300/J075v27n03_01

Sran, S. K., & Borrero, J. C. (2010). Assessing the value of choice in a token system. *Journal of Applied Behavior Analysis, 43*, 553–557.

St. Peter Pipkin, C. (2015). Six reasons why applied behavior analysts should know about resurgence. *Mexican Journal of Behavior Analysis*, *41*, 252–268.

St. Peter Pipkin, C., Vollmer, T. R., & Sloman, K. N. (2010). Effects of treatment integrity failures during differential reinforcement of alternative behavior: A translational model. *Journal of Applied Behavior Analysis*, *43*, 47–70.

St. Peter, C. C., Byrd, J. D., Pence, S. T., & Foreman, A. P. (2016). Effects of treatment-integrity failures on a response-cost procedure. *Journal of Applied Behavior Analysis*, *49*, 308–328. doi:10.1002/jaba.291

Staats, A. W. (1996). *Behavior and personality: Psychological behaviorism*. New York: Springer Publishing.

Staats, A. W., & Staats, C. K. (1963). *Complex human behavior: A systematic extension of learning principles*. New York: Holt, Rinehart and Winston.

Staats, H., van Leeuwen, E., & Wit, A. (2000). A longitudinal study of informational interventions to save energy in an office building. *Journal of Applied Behavior Analysis*, *33*, 101–104.

Staddon, J. E. R. (1977). Schedule-induced behavior. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 125–152). Upper Saddle River, NJ: Prentice Hall.

Stafford, M. W., Sundberg, M. L., & Braam, S. J. (1988). A preliminary investigation of the consequences that define the mand and the tact. *The Analysis of Verbal Behavior*, *6*, 61–71.

Stafford, T. (2015, November 16). *No more Type I/II error confusion*. Retrieved January 3, 2019, from https://mindhacks.com/2015/11/16/no-more-type-iii-error-confusion/

Stage, S. A., & Quiroz, D. R. (1997). A meta-analysis of interventions to decrease disruptive classroom behavior in public education settings. *School Psychology Review*, *26*, 333–368.

Stecker, P. M., Fuchs, L. S., & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of research. *Psychology in the Schools*, *42*(8), 795–819.

Steege, M. W., Wacker, D. P., Cigrand, K. C., Berg, W. K., Novak, C. G., Reimers, T. M., . . . DeRaad, A. (1990). Use of negative reinforcement in the treatment of self-injurious behavior. *Journal of Applied Behavior Analysis*, *23*, 459–467.

Steingrimsdottir, H. S., & Arntzen, E. (2011). Using conditional discrimination procedures to study remembering in an Alzheimer's patient. *Behavioral Interventions*, *26*, 179–192.

Stephens, T. M. (1976). *Directive teaching of children with learning and behavioral handicaps*. Columbus, OH: Merrill, 1976.

Steuart, W. (1993). Effectiveness of arousal and arousal plus overcorrection to reduce nocturnal bruxism. *Journal of Behavior Therapy & Experimental Psychiatry*, *24*, 181–185.

Stevens, K. B., & Schuster, J. W. (1987). Effects of a constant time delay procedure on the written spelling performance of a learning disabled student. *Learning Disability Quarterly*, *10*, 9–16.

Stevenson, H. C., & Fantuzzo, J. W. (1984). Application of the "generalization map" to a self-control intervention with school-aged children. *Journal of Applied Behavior Analysis*, *17*, 203–212.

Stevenson, M. T., Ghezzi, P. M., & Valenton K. G. (2016). FCT and delay fading for elopement with a child with autism. *Behavior Analysis in Practice*, *9*, 169–173.

Stewart, C. A., & Singh, N. N. (1986). Overcorrection of spelling deficits in mentally retarded persons. *Behavior Modification*, *10*, 355–365.

Stewart, I., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (2002). A functional-analytic model of analogy: A relational frame analysis. *Journal of the Experimental Analysis of Behavior*, *78*, 375–396.

Stewart, I., McElwee, J., & Ming, S. (2013). Language generativity, response generalization and derived relational responding. *The Analysis of Verbal Behavior*, *29*, 137–155.

Stewart, K. K., Carr, J. E., Brandt, C. W., & McHenry, M. M. (2007). An evaluation of the conservative dual-criterion method for teaching university students to visually inspect A-B design graphs. *Journal of Applied Behavior Analysis*, *40*, 713–718.

Stilitz, I. (2009). A token economy of the early 19th century. *Journal of Applied Behavior Analysis*, *42*, 925–926.

Stocco, C. S., Thompson, R. H., & Hart, J. M. (2014). Teaching tacting of private events based on public accompaniments: Effects of contingencies, audience control, and stimulus complexity. *The Analysis of Verbal Behavior*, *30*, 1–19. doi.org/10.1007/s40616-014-0006-y

Stock, R. A., Schulze, K. A., & Mirenda, P. (2008). A comparison of stimulus-stimulus pairing, standard echoic training, and control procedures, of the vocal behavior of children with autism. *The Analysis of Verbal Behavior*, *24*, 123–133.

Stockard, J., Wood, T. W., Coughin, C., & Khoury, C. R. (2018). The effectiveness of Direct Instruction curricula: A meta-analysis of a half century of research. *Review of Educational Research*. Advance online publication. doi:10.3102/0034654317751919

Stoddard, L., Brown, J., Hurlbert, B., Manoli, C., & McIlvane, W. J. (1989). Teaching money skills through stimulus class formation, exclusion, and component matching methods: Three case studies. *Research in Developmental Disabilities*, *10*, 413–439.

Stokes, J. V., Luiselli, J. K., & Reed, D. D. (2010). A behavioral intervention for teaching tackling skills to high school football athletes. *Journal of Applied Behavior Analysis*, *43*, 509–512. doi:0.1901/jaba.2010.43-509

Stokes, J. V., Luiselli, J. K., Reed, D. D., & Fleming, R. K. (2010). Behavioral coaching to improve offensive line pass-blocking skills of high school football athletes. *Journal of Applied Behavior Analysis*, *43*, 463–472.

Stokes, T. (2003). A genealogy of applied behavior analysis. In K. S. Budd & T. Stokes (Eds.), *A small matter of proof: The legacy of Donald M. Baer* (pp. 257–272). Reno, NV: Context Press.

Stokes, T. F., & Baer, D. M. (1976). Preschool peers as mutual generalization-facilitating agents. *Behavior Therapy*, *7*, 599–610.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, *10*, 349–367. doi:10.1901/jaba.1977.10-349

Stokes, T. F., & Osnes, P. G. (1982). Programming the generalization of children's social behavior. In P. S. Strain, M. J. Guralnick, & H. M. Walker (Eds.), *Children's social behavior: Development, assessment, and modification* (pp. 407–443). Orlando, FL: Academic Press.

Stokes, T. F., & Osnes, P. G. (1989). An operant pursuit of generalization. *Behavior Therapy*, *20*, 337–355.

Stokes, T. F., Baer, D. M., & Jackson, R. L. (1974). Programming the generalization of a greeting response in four retarded children. *Journal of Applied Behavior Analysis*, *7*, 599–610.

Stokes, T. F., Fowler, S. A., & Baer, D. M. (1978). Training preschool children to recruit natural communities of reinforcement. *Journal of Applied Behavior Analysis*, *11*, 285–303.

Stokoe, W. C., Casterline, D. C., & Croneberg. C. G. (1965). *A dictionary of American sign languages on linguistic principles*. Washington, DC: Gallaudet College Press.

Stolz, K. E., Itoi, M., Konrad, M., & Alber-Morgan, S. R. (2008). Effects of self-graphing on written expression of fourth grade students with high-incidence disabilities. *Journal of Behavioral Education*, *17*, 172–186. doi 10.1007/s10864-007-9055-9

Stolz, S. B. (1978). *Ethical issues in behavior modification*. San Francisco: Jossey-Bass.

Strachey, L. (1921). *Queen Victoria*. New York: Harcourt, Brace.

Strain, P. S., & Joseph, G. E. (2004). A not so good job with "Good job." *Journal of Positive Behavior Interventions*, *6*(1), 55–59.

Strain, P. S., Shores, R. E., & Kerr, M. M. (1976). An experimental analysis of "spillover" effects on the social interaction of behaviorally handicapped preschool children. *Journal of Applied Behavior Analysis*, *9*, 31–40.

Striefel, S. (1974). *Behavior modification: Teaching a child to imitate*. Austin, TX: PRO-ED.

Stromer, R. (2000). Integrating basic and applied research and the utility of Lattal and Perone's *Handbook of Research Methods in Human Operant Behavior. Journal of Applied Behavior Analysis*, *33*, 119–136. doi:10.1901/jaba.2000.33-119

Stromer, R., & Mackay, H. A. (1992). Spelling and emergent picture-printed word relations established with delayed identity matching to complex samples. *Journal of Applied Behavior Analysis*, *25*, 893–904.

Stromer, R., McComas, J. J., & Rehfeldt, R. A. (2000). Designing interventions that include delayed reinforcement: Implications of recent laboratory research. *Journal of Applied Behavior Analysis*, *33*, 359–371.

Stromer, R., McComas, J. J., and Rehfeldt, R. A. (2000). Effects of intermittent punishment on self-injurious behavior: An evaluation of schedule thinning. *Journal of Applied Behavior Analysis*, *30*, 187–201.

Suess, A. N., Romani, P. W., Wacker, D. P., Dyson, S. M., Kuhle, J. L., Lee, J. F., . . . Waldron, D. B. (2014). Evaluating the treatment fidelity of parents who conduct in-home functional communication training with coaching via telehealth. *Journal of Behavioral Education*, *23*, 34–59.

Sugai, G. M., & Tindal, G. A. (1993). *Effective school consultation: An interactive approach*. Pacific Grove, CA: Brooks/Cole.

Sulzer-Azaroff, B., & Mayer, G. R. (1977). *Applying behavior-analysis procedures with children and youth*. New York: Holt, Rinehart and Winston.

Sulzer-Azaroff, B., Mayer, G. R., & Wallace, M. (2013). *Behavior analysis for lasting change* (3rd ed.). Cornwall-on-Hudson, NY: Sloan Publishing.

Sundberg, C. T., & Sundberg, M. L. (1990). Comparing topography-based verbal behavior with stimulus selection-based verbal behavior. *The Analysis of Verbal Behavior*, *8*, 31–41.

Sundberg, C. T., Sundberg, M. L., & Michael, J. (2018). Covert verbal mediation and arbitrary matching to sample. *The Journal of the Experimental Analysis of Behavior*, *109*, 600–623. doi.org/10.1002/jeab.434

Sundberg, M. L. (1983). Language. In J. L. Matson & S. E. Breuning (Eds.), *Assessing the mentally retarded* (pp. 285–310). New York: Grune & Stratton.

Sundberg, M. L. (1993a). Selecting a response form for nonverbal persons: Facilitated communication, pointing systems, or sign language? *The Analysis of Verbal Behavior*, *11*, 99–116.

Sundberg, M. L. (1993b). The application of establishing operations. *The Behavior Analyst*, *16*, 211–214.

Sundberg, M. L. (2013). Thirty points about motivation from Skinner's book *Verbal Behavior*. *The Analysis of Verbal Behavior*, *27*, 13–40.

Sundberg, M. L. (2014). *The verbal behavior milestones assessment and placement program: The VB-MAPP* (2nd ed.). Concord, CA: AVB Press.

Sundberg, M. L. (2015). The most important verbal operant. *VB-NEWS*, *14*(2), 3–5.

Sundberg, M. L. (2016a). Verbal stimulus control and the intraverbal relation. *The Analysis of Verbal Behavior*, *32*, 107–124.

Sundberg, M. L. (2016b). The value of a behavioral analysis of language for autism treatment. In R. G. Romanczyk & J. McEachin (Eds.), *Comprehensive models of autism spectrum disorder treatment: Points of divergence and convergence*. New York: Springer.

Sundberg, M. L., & Michael, J. (2001). The benefits of Skinner's analysis of verbal behavior for teaching children with autism. *Behavior Modification*, *25*, 698–724.

Sundberg, M. L., & Partington, J. W. (1998). *Teaching language to children with autism or other developmental disabilities*. Pleasant Hill, CA: Behavior Analysts, Inc.

Sundberg M. L., & Sundberg, C. A. (2011). Intraverbal behavior and verbal conditional discriminations in typically developing children and children with autism. *The Analysis of Verbal Behavior*, *27*, 23–43.

Sundberg, M. L., Loeb, M., Hale, L., & Eigenheer, P. (2002). Contriving establishing operations to teach mands for information. *The Analysis of Verbal Behavior*, *18*, 14–28.

Sundberg, M. L., Michael, J., Partington, J. W., & Sundberg, C. A. (1996). The role of automatic reinforcement in early language acquisition. *The Analysis of Verbal Behavior*, *13*, 21–37.

Sundberg, M. L., Ray, D. A., Braam, S. J., Stafford, M. W., Rueber, T. M., & Braam, C. (1979). A manual for the use of B. F. Skinner's analysis of verbal behavior for language assessment and programming. *Western Michigan University Behavioral Monograph*, *#9*. Retrieved from http://www.marksundberg.com/files/8-WMU_Monograpgh_9.pdf

Sundberg, M. L., San Juan, B., Dawdy, M., & Arguelles, M. (1990). The acquisition of tacts, mands, and intraverbals by individuals with traumatic brain injury. *The Analysis of Verbal Behavior*, *8*, 83–99.

Superintendent of Belchertown State School v. Saikewicz, 373 Mass. 728 (1977) 370 N. E. 2d 417.

Sutherland, K. S., Wehby, J. H., & Yoder, P. J. (2002). An examination of the relation between teacher praise and students with emotional/behavioral disorders' opportunities to respond to academic requests. *Journal of Emotional and Behavioral Disorders*, *10*, 5–13.

Swaminathan, H., Rogers, H. J., & Horner, R. (2014). An effect size measure and Bayesian analysis of single case designs. *Journal of School Psychology*, *52*, 213–230. doi.org/10.1016/j.jsp.2013.12.002

Swanson, H. L., & Sachse-Lee, C. (2000). A meta-analysis of single-subject-design intervention research for students with LD. *Journal of Learning Disabilities*, *38*, 114–136.

Swanson, N. L., Leu, A., Abrahamson, J. & Wallet, B. (2014). Genetically engineered crops, glyphosate and the deterioration of health in the United States of America. *Journal of Organic Systems*, *9*(2), 6–37.

Symons, F. J., Hoch, J., Dahl, N. A., & McComas, J. J. (2003). Sequential and matching analyses of self-injurious behavior: A case of overmatching in the natural environment. *Journal of Applied Behavior Analysis*, *36*, 267–270.

Symons, F. J., McDonald, L. M., & Wehby, J. H. (1998). Functional assessment and teacher collected data. *Education and Treatment of Children*, *21*, 135–159.

Szabo, T., Bockaris, H., White, R., Tarbox, J., Stewart, C., & Ebert, B. (2019). *Microaggression, intimate partner gender-based violence, and behavioral flexibility training in Sierra Leonean couples*. Manuscript under review.

Tang, J., Kennedy, C. H., Koppekin, A., & Caruso, M. (2002). Functional analysis of stereotypical ear covering in a child with autism. *Journal of Applied Behavior Analysis*, *35*, 95–98.

Tankersley, M. (1995). A group-oriented contingency management program: A review of research on the Good Behavior Game and implications for teachers. *Preventing School Failure*, *40*, 19–24.

Tanno, T., Silberberg, A., & Sakagami, T. (2009). Single-sample discrimination of different schedules' reinforced interresponse times. *Journal of the Experimental Analysis of Behavior*, *37*, 157–167.

Tanol, G., Johnson, L., McComas, J., & Cote, E. (2010). Responding to rule violations or rule following: A comparison of two versions of the Good Behavior Game with kindergarten students. *Journal of School Psychology*, *48*, 337–355.

Tapp, J. T., & Walden, T. A. (2000). A system for collecting and analysis of observational data from videotape. In T. Thompson, D. Felce, & F. J. Symons (Eds.), *Behavioral observation: Technology and applications in developmental disabilities* (pp. 61–70). Baltimore: Paul H. Brookes.

Tapp, J. T., & Wehby, J. H. (2000). Observational software for laptop computers and optical bar code readers. In T. Thompson, D. Felce, & F. J. Symons (Eds.), *Behavioral observation: Technology and applications in developmental disabilities* (pp. 71–81). Baltimore: Paul H. Brookes.

Tapp, J. T., Wehby, J. H., & Ellis, D. M. (1995). A multiple option observation system for experimental studies: MOOSES. *Behavior Research Methods Instruments & Computers*, *27*, 25–31.

Tarasoff v. Regents of University of California, 551 P. 2d 334 – Cal: Supreme Court 1976. https://scholar.google.com/scholar_case?case=263231934673470561

Tarbox, J., Madrid, W., Aguilar, B., Jacobo, W., & Schiff, A. (2009). Use of chaining to increase complexity of echoics in children with autism. *Journal of Applied Behavior Analysis*, *42*, 901–906.

Tarbox, R. S. F., Wallace, M. D., & Williams, L. (2003). Assessment and treatment of elopement: A replication and extension. *Journal of Applied Behavior Analysis*, *36*, 239–244.

Tarbox, R. S. F., Williams, W. L., & Friman, P. C. (2004). Extended diaper wearing: Effects on continence in and out of the diaper. *Journal of Applied Behavior Analysis*, *37*, 101–105.

Tate, R., Perdices, M., Rosenkoetter, U., McDonald, S., Togher, L., Shadish, W., . . . for the SCRIBE Group. (2016a). The single-case reporting guideline in BEhavioural interventions (SCRIBE) 2016 statement. *Archives of Scientific Psychology*, *4*, 1-9.

Tate, R., Perdices, M., Rosenkoetter, U., McDonald, S., Togher, L., Shadish, W., . . . for the SCRIBE Group. (2016b). The single-case reporting guideline in BEhavioural interventions (SCRIBE) 2016: Explanation and elaboration. *Archives of Scientific Psychology*, *4*, 10–31.

Tawney, J., & Gast, D. (1984). *Single subject research in special education*. Columbus, OH: Charles E. Merrill.

Taylor, B. A. (2013). "Do this, but don't do that": Teaching children with autism to learn by observation. In W. L. Heward (Ed.), *Exceptional children: An introduction to special education* (10th ed.) (pp. 240–242). Upper Saddle River, NJ: Pearson.

Taylor, B. A., & DeQuinzio, J. A. (2012). Observational learning and children with autism. *Behavior Modification*, *36*, 341–360.

Taylor, B. A., & Hoch, H. (2008). Teaching children with autism to respond to and initiate bids for joint attention. *Journal of Applied Behavior Analysis*, *41*, 377–391.

Taylor, B. A., DeQuinzio, J. A., & Stine, J. (2012). Increasing observational learning of children with autism: A preliminary analysis. *Journal of Applied Behavior Analysis*, *45*, 815–820. doi:10.1901/ jaba.2012.45-815

Taylor, B. A., Hoch, H., Potter, B., Rodriguez, A., Spinnato, D., & Kalaigian, M. (2005). Manipulating establishing operations to promote initiations toward peers in children with autism. *Research in Developmental Disabilities, 26*, 385–392.

Taylor, I., & O'Reilly, M. F. (2000). Generalization of supermarket shopping skills for individuals with mild intellectual disabilities using stimulus equivalence training. *The Psychological Record, 50*, 49–62.

Temple, C. H., Ogle, D., Crawford, A. N., & Freppon, P. (2018). *All children read: Teaching for literacy in today's diverse classrooms* (5th ed.). Upper Saddle River, NJ: Pearson.

Terrace, H. S. (1963a). Discrimination learning with and without "errors." *Journal of the Experimental Analysis of Behavior, 6*, 1–27.

Terrace, H. S. (1963b). Errorless transfer of a discrimination across two continua. *Journal of the Experimental Analysis of Behavior, 6*, 223–232.

Terris, W., & Barnes, M. (1969). Learned resistance to punishment and subsequent responsiveness to the same and novel punishers. *Psychonomic Science, 15*, 49–50.

Tertinger, D. A., Greene, B. F., & Lutzker, J. R. (1984). Home safety: Development and validation of one component of an ecobehavioral treatment program for abused and neglected children. *Journal of Applied Behavior Analysis, 17*, 159–174.

Test, D. W., & Heward, W. L. (1983). Teaching road signs and traffic laws to learning disabled students. *Science Education, 64*, 129–139.

Test, D. W., & Heward, W. L. (1984). Accuracy of momentary time sampling: A comparison of fixed and variable-interval observation schedules. In W. L. Heward, T. E. Heron, D. S. Hill, and J. Trap-Porter (Eds.), *Focus on behavior analysis in education* (pp. 177–194). Columbus, OH: Charles E. Merrill.

Test, D. W., Spooner, F., Keul, P. K., & Grossi, T. (1990). Teaching adolescents with severe disability to use the public telephone. *Behavior Modification, 14*, 157–171.

Thaler, R. H. (2012, February 13). Making good citizenship fun. *The New York Times*, p. A25. [Available at http://www.nytimes.com/2012/02/14/opinion/making-good-citizenship-fun.html?]

Theobold, D. E., & Paul, G. L. (1976). Reinforcing value of praise for chronic mental patients as a function of historical pairing with tangible reinforcers. *Behavior Therapy, 7*, 192–197.

Theodore, L. A., Bray, M. A., Kehle, T. J., & Jenson, W. R. (2001). Randomization of group contingencies and reinforcers to reduce classroom disruptive behavior. *Journal of School Psychology, 39*, 267–277.

Thomason-Sassi, J. L., Iwata, B. A., Neidert, P. L., & Roscoe, E. M. (2011). Response latency as an index of response strength during functional analyses of problem behavior. *Journal of Applied Behavior Analysis, 44*, 51–67. doi:10.1901/jaba.2011.44-51

Thompson, L. G. (2010). Climate change: The evidence and our options. *The Behavior Analyst, 33*, 153–170.

Thompson, R. H., & Borrero, J. C. (2011). Direct observation. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 191–205). New York: Guilford Press.

Thompson, R. H., & Iwata, B. A. (2000). Response acquisition under direct and indirect contingencies of reinforcement. *Journal of Applied Behavior Analysis, 33*, 1–11.

Thompson, R. H., & Iwata, B. A. (2001). A descriptive analysis of social consequences following problem behavior. *Journal of Applied Behavior Analysis, 34*, 169–178.

Thompson, R. H., & Iwata, B. A. (2005). A review of reinforcement control procedures. *Journal of Applied Behavior Analysis, 38*, 257–278.

Thompson, R. H., Bruzek, J. L., & Cotnoir-Bichelman, N. M. (2011). The role of negative reinforcement in infant caregiving: An experimental simulation. *Journal of Applied Behavior Analysis, 44*, 295-304.

Thompson, R. H., Iwata, B. A., Conners, J., & Roscoe, E. M. (1999). Effects of reinforcement for alternative behavior during punishment of self-injury. *Journal of Applied Behavior Analysis, 32*, 317–328.

Thompson, T., & Hackenberg, T. D. (2009). Introduction: Translational science lectures. *The Behavior Analyst, 32*, 269–271.

Thompson, T., Symons, F. J., & Felce, D. (2000). Principles of behavioral observation. In T. Thompson, F. J. Symons, & D. Felce (Eds.), *Behavioral observation: Technology and applications in developmental disabilities* (pp. 3–16). Baltimore: Paul H. Brookes.

Thomson, C., Holmber, M., & Baer, D. M. (1974). A brief report on a comparison of time sampling procedures. *Journal of Applied Behavior Analysis, 7*, 623–626.

Thoresen, C. E., & Mahoney, M. J. (1974). *Behavioral self-control*. New York: Holt, Rinehart & Winston.

Thurber, J. (1999). *The Thurber carnival* (Perennial Classics Edition, reissue). New York: HarperCollins.

Tiger, J. H., & Hanley, G. P. (2004). Developing stimulus control of preschooler mands: An analysis of schedule-correlated and contingency-specifying stimuli. *Journal of Applied Behavior Analysis, 37*, 517–521.

Tiger, J. H., & Hanley, G. P. (2005). An example of discovery research involving the transfer of stimulus control. *Journal of Applied Behavior Analysis, 38*, 499–509.

Tiger, J. H., & Kliebert, M. L. (2011). Stimulus preference assessment. In J. K. Luiselli (Ed.), *Teaching and behavior support for children and adults with autism spectrum disorder: A practitioner's guide* (pp. 30–37). New York: Oxford.

Tiger, J. H., Fisher, W. W., & Bouxsein, K. J. (2009). Therapist- and self-monitored DRO contingencies as a treatment for the self-injurious skin picking of a young man with Asperger syndrome. *Journal of Applied Behavior Analysis, 42*, 315–319. http://doi.org/10.1901/jaba.2009.42-315

Tiger, J. H., Hanley, G. P., & Heal, N. A. (2006). The effectiveness of and preschoolers' preferences for variations of multiple-schedule arrangements. *Journal of Applied Behavior Analysis, 39*, 475–488.

Tiger, J. H., Miller, S. J., Mevers, J. L., Mintz, J. C., Scheithauer, M. C., & Alvarez, J. (2013). On the representativeness of behavior observation samples in classrooms. *Journal of Applied Behavior Analysis*, 46, 424–435. doi:10.1002/jaba.39

Timberlake, W., & Allison, J. (1974). Response deprivation: An empirical approach to instrumental performance. *Psychological Review, 81*, 146–164.

Tincani, M. (2004). Comparing the picture exchange communication system and sign language training for children with autism. *Focus on Autism and Other Developmental Disabilities, 19*, 152–163.

Tincani, M. J., Castrogiavanni, A., & Axelrod, S. (1999). A comparison of the brief versus traditional functional analyses. *Research in Developmental Disabilities, 20*, 327–338.

Tingstrom, D. H., Sterling-Tuner, H. E., & Wilczynski, S. M. (2006), The Good Behavior Game: 1969–2002. *Behavior Modification, 30*(2), 225–253.

Toates, F. (2009). *Burrhus F. Skinner: Shaper of behaviour*. London, UK: Palgrave Macmillan.

Todd, A., Campbell, A., Meyer, G., & Horner, R. (2008). Evaluation of a targeted group intervention in elementary students: The check-in/check-out program. *Journal of Positive Behavior Interventions, 10*, 46–55.

Todd, A. W., Horner, R. W., & Sugai, G. (1999). Self-monitoring and self-recruited praise: Effects on problem behavior, academic engagement, and work completion in a typical classroom. *Journal of Positive Behavior Interventions, 1*, 66–76, 122.

Todd, J. T. (2012). The moral obligation to be empirical: Comments on Boynton's "Facilitated Communication—what harm it can do: Confessions of a former facilitator." *Evidence-Based Communication Assessment and Intervention, 6*(1), 36–57.

Todd, J. T., & Morris, E. K. (1983). Misconception and miseducation: Presentations of radical behaviorism in psychology textbooks. *The Behavior Analyst, 6*, 153–160.

Todd, J. T., & Morris, E. K. (1992). Case histories in the great power of steady misrepresentation. *American Psychologist, 47*, 1441–1453.

Todd, J. T., & Morris, E. K. (1993). Change and be ready to change again. *American Psychologist, 48*, 1158–1159.

Todes, D. P. (2014). *Ivan Pavlov: A Russian life in science*. New York: Oxford University Press.

Tommis, Y., & Fazey, D. M. A. (1999). The acquisition of the pitch element of music literacy skills by 3-4-year-old pre-school children: A comparison of two methods. *Psychology of Music, 27*, 230–244.

Toogood, S. (2012). Using contingency diagrams in the functional assessment of challenging behaviour. *International Journal of Positive Behavioural Support, 2*(1), 3–10.

Torelli, J. N., Lloyd, B. P., Diekman, C. A., & Wehby, J. H. (2017). Teaching stimulus control via class-wide multiple schedules of reinforcement in public elementary school classrooms. *Journal of Positive Behavior Interventions, 19*, 14–25.

Torelli, J. N., Lloyd, B. P., Diekman, C. A., & Wehby, J. H. (2016). Teaching stimulus control via class-wide multiple schedules of reinforcement in public elementary school classrooms. *Journal of Positive Behavior Interventions, 18*, 1–12. doi:10.1177/1098300716632878

Torneke, N. (2010). *Learning RFT: An introduction to relational frame theory and its clinical application*. Oakland, CA: New Harbinger Publications.

Touchette, P. (1971). Transfer of stimulus control: Measuring the moment of transfer. *Journal of the Experimental Analysis of Behavior, 6*, 223–232.

Touchette, P. E., MacDonald, R. F., & Langer, S. N. (1985). A scatter plot for identifying stimulus control of problem behavior. *Journal of Applied Behavior Analysis, 18*, 343–351.

Tourinho, E. Z., & Neno, S. (2003). Effectiveness as truth criterion in behavior analysis. *Behavior and Philosophy, 31*, 63–80.

Toussaint, K. A., & Tiger, J. H. (2010). Teaching early braille literacy skills within a stimulus equivalence paradigm to children with degenerative visual impairments. *Journal of Applied Behavior Analysis, 43*, 181–194.

Towery, D., Parsons, M. B., & Reid, D. H. (2014). Increasing independence within adult services: A program for reducing staff completion of daily routines for consumers with developmental disabilities. *Behavior Analysis in Practice, 7*, 61–69. doi 10.1007/s40617-014-0013-4

Trahan, M. A., Donaldson, J. M., McNabney, M. K., & Kahng, S. (2014). Training and maintenance of a picture-based communication response in older adults with dementia. *Journal of Applied Behavior Analysis, 47*, 404–409. doi:10.1002/jaba.111

Trap, J. J., Milner-Davis, P., Joseph, S., & Cooper, J. O. (1978). The effects of feedback and consequences on transitional cursive letter formation. *Journal of Applied Behavior Analysis, 11*, 381–393.

Trask-Tyler, S. A., Grossi, T. A., & Heward, W. L. (1994). Teaching young adults with developmental disabilities and visual impairments to use tape-recorded recipes: Acquisition, generalization, and maintenance of cooking skills. *Journal of Behavioral Education, 4*, 283–311.

Travis, R., & Sturmey, P. (2010). Functional analysis and treatment of the delusional statements of a man with multiple disabilities: A four-year follow-up. *Journal of Applied Behavior Analysis, 43*, 745–749.

Tu, J. C. (2006). The role of joint control in the manded selection responses of both vocal and non-vocal children with autism. *The Analysis of Verbal Behavior, 22*, 191–207.

Tucker, D. J., & Berry, G. W. (1980). Teaching severely multihandicapped students to put on their own hearing aids. *Journal of Applied Behavior Analysis, 13*, 65–75.

Tufte, E. R. (2001). *The visual display of quantitative information* (2nd ed.). Chesire, CT: Graphics Press.

Tufte, E. R. (2017). *Baseline for amount* scale. Retrieved March 26, 2017, at https://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=00003q

Tung, S. B., Donaldson, J. M., & Kahng, S. W. (2017).The effects of preference assessment type on problem behavior. *Journal of Applied Behavior Analysis, 50*, 861–866. doi:10.1002/jaba.414

Turkewitz, H., O'Leary, K. D., & Ironsmith, M. (1975). Generalization and maintenance of appropriate behavior through self-control. *Journal of Consulting and Clinical Psychology, 43*, 577–583.

Turnbull, H. R., III., & Turnbull, A. P. (1998). *Free appropriate public education: The law and children with disabilities*. Denver, CO: Love.

Turner, L. B., Fisher, A. J., & Luiselli, J. K. (2016). Towards a competency-based, ethical, and socially valid approach to supervision of applied behavior analytic trainees. *Behavior Analysis in Practice, 9*, 287–298. doi:10.1007/s40617-016-0121-4

Tustin, R. D. (1994). Preference for reinforcers under varying schedule arrangements: A behavioral economic analysis. *Journal of Applied Behavior Analysis, 28*, 61–71.

Twohig, M. P., & Woods, D. W. (2001). Evaluating the duration of the competing response in habit reversal: A parametric analysis. *Journal of Applied Behavior Analysis, 34*, 517–520.

Twyman, J., Johnson, H., Buie, J., & Nelson, C. M. (1994). The use of a warning procedure to signal a more intrusive timeout contingency. *Behavioral Disorders, 19*(4), 243–253.

Twyman, J. S. (2010). TerraKids. *The Behavior Analyst, 33*, 193–196.

Twyman, J. S. (2011). Emerging technologies and behavioural cusps: A new era for behaviour analysis. *European Journal of Behaviour Analysis, 12*, 461–482.

Twyman, J. S. (2013). Behavior analysis in education. In F. K. McSweeney & E. S. Murphy (Eds.), *The Wiley-Blackwell handbook of operant and classical conditioning* (pp. 533–558). Oxford, UK: John Wiley & Sons.

Twyman, J. S., Layng, T. V., Stikeleather, G., & Hobbins, K. A. (2005). A non-linear approach to curriculum design: The role of behavior analysis in building an effective reading program. In W. L. Heward, T. E. Heron, N. A. Neef, S. M. Peterson, D. M. Sainato, G. Cartledge, R. Gardner, III, L. D. Peterson, S. B. Hersh, & J. C. Dardig (Eds.), *Focus on behavior analysis in education: Achievements, challenges, and opportunities* (pp. 55–67). Upper Saddle River, NJ: Pearson.

Tyner, B. C., & Fienup, D. M. (2015). A comparison of video modeling, text-based instruction, and no instruction for creating multiple baseline graphs in Microsoft Excel. *Journal of Applied Behavior Analysis, 48*, 701–706.

Tyner, S., Brewer, A., Helman, M., Leon, Y., Pritchard, J., & Schlund, M. (2016). Nice doggie! Contact desensitization plus reinforcement decreases dog phobias for children with autism. *Behavior Analysis in Practice, 9*, 54–57.

Ulman, J. D., & Sulzer-Azaroff, B. (1975). Multi-element baseline design in educational research. In E. Ramp & G. Semb (Eds.), *Behavior analysis: Areas of research and application* (pp. 371–391). Upper Saddle River, NJ: Prentice Hall.

Ulrich, R. E., & Azrin, N. H. (1962). Reflexive fighting in response to aversive stimulation. *Journal of the Experimental Analysis of Behavior, 5*, 511–520.

Ulrich, R. E., Stachnik, T., & Mabry, J. (Eds.). (1974). *Control of human behavior (Vol. 3), Behavior modification in education*. Glenview, IL: Scott, Foresman and Company.

Ulrich, R. E., Wolff, P. C., & Azrin, N. H. (1962). Shock as an elicitor of intra and interspecies fighting behavior. *Animal Behavior, 12*, 14–15.

Upper, D. (1974). The unsuccessful self-treatment of a case of "writer's block." *Journal of Applied Behavior Analysis, 7*, 497.

U.S. Department of Education. (2003). *Identifying and implementing educational practices supported by rigorous evidence: A user friendly guide*. Washington, DC: Institute of Education Sciences National Center for Education Evaluation and Regional Assistance.

Uttal, W. R. (2000). *The war between mentalism and behaviorism: On the accessibility of mental processes*. Mahwah, NJ: Erlbaum.

Valbuena, D., Miltenberger, R., & Solley, E. (2015). Evaluating an Internet-based program and a behavioral coach for increasing physical activity. *Behavior Analysis Research and Practice, 15*, 122–138.

Valdivia-Salas, S., Luciano, C., Gutierrez-Martinez, O., & Visdomine, C. (2009). Establishing empathy. In R. A. Rehfeldt & Y. Barnes-Holmes (Eds.), *Derived relational responding: Applications for learners with autism and other developmental disabilities: A progressive guide to change* (pp. 301–311). New Oakland, CA: Harbinger Publications.

Valentino, A. L., Sellers. T. P., & LeBlanc, L. A. (2016). The benefits of group supervision and a recommended structure for implementation. *Behavior Analysis in Practice, 9*, 320–328. doi10.1007/s40617-016-0138-8

Valk, J. E. (2003). *The effects of embedded instruction within the context of a small group on the acquisition of imitation skills of young children with disabilities*. Unpublished doctoral dissertation, The Ohio State University, Columbus, OH.

Van Camp, C. M., Borrero, J. C., & Vollmer, T. R. (2003). The family safety/applied behavior analysis initiative: An introduction and overview. *The Behavior Analyst Today, 3*, 389–404.

Van Camp, C. M., Lerman, D. C., Kelley, M. E., Contrucci, S. A., & Vorndran, C. M. (2000). Variable-time reinforcement schedules in the treatment of socially maintained problem behavior. *Journal of Applied Behavior Analysis, 33*, 545–557.

van den Pol, R. A., Iwata, B. A., Ivancic, M. T., Page, T. J., Neef, N. A., & Whitley, F. P. (1981). Teaching the handicapped to eat in public places: Acquisition, generalization and maintenance of restaurant skills. *Journal of Applied Behavior Analysis, 14*, 61–69.

Van der Gucht, K., Griffith, J. W., Hellemans, R., Bockstaele, M., Pascal-Claes, F., & Raes, F. (2017). Acceptance and commitment therapy (ACT) for adolescents: Outcomes of a large-sample, school-based, cluster-randomized controlled trial. *Mindfulness, 8*, 408–416.

van Harran, F. (2015). Automatic negative reinforcement: Its possible role in problem behavior and treatment implications. *Behavior Analysis: Research and Practice, 15*, 161–170.

Van Houten, R. (1979). Social validation: The evolution of standards of competency for target behaviors. *Journal of Applied Behavior Analysis*, *12*, 581–591.

Van Houten, R. (1979). The performance feedback system. *Child Behavior Therapy*, 1(3), 219–236. doi:10.1300/J473v01n03_01

Van Houten, R. (1993). Use of wrist weights to reduce self-injury maintained by sensory reinforcement. *Journal of Applied Behavior Analysis*, *26*, 197–203.

Van Houten, R., & Doleys, D. M. (1983). Are social reprimands effective? In S. Axelrod & J. Apsche (Eds.), *The effects of punishment on human behavior* (pp. 45–70). New York: Academic Press.

Van Houten, R., & Malenfant, J. E. L. (2004). Effects of a driver enforcement program on yielding to pedestrians. *Journal of Applied Behavior Analysis*, *37*, 351–363.

Van Houten, R., & Nau, P. A. (1980). A comparison of the effects of fixed and variable ratio schedules of reinforcement on the behavior of deaf children. *Journal of Applied Behavior Analysis*, *13*, 13–21.

Van Houten, R., & Rolider, A. (1984). The use of response prevention to eliminate nocturnal thumbsucking. *Journal of Applied Behavior Analysis*, *17*, 509–520.

Van Houten, R., & Rolider, A. (1988). Recreating the scene: An effective way to provide delayed punishment for inappropriate motor behavior. *Journal of Applied Behavior Analysis*, *21*, 187–192.

Van Houten, R., & Rolider, A. (1990). The role of reinforcement in reducing inappropriate behavior: Some myths and misconceptions. In A. C. Repp & N. N. Singh (Eds.), *Perspectives on the use of nonaversive and aversive interventions for persons with developmental disabilities* (pp. 119–127). Sycamore, IL: Sycamore.

Van Houten, R., Axelrod, S., Bailey, J. S., Favell, J. E., Foxx, R. M., Iwata, B. A., & Lovaas, O. I. (1988). The right to effective behavioral treatment. *Journal of Applied Behavior Analysis*, *21*, 381–384.

Van Houten, R., Axelrod, S., Bailey, J. S., Favell, J. E., Foxx, R. M., Iwata, B. A., & Lovaas, O. I. (1988). The right to effective behavioral treatment. *The Behavior Analyst*, *11*, 111–114.

Van Houten, R., Malenfant, J. E., Reagan, I., Sifrit, K., Compton, R., & Tenenbaum, J. (2010). Increasing seat belt use in service vehicle drivers with a gearshift delay. *Journal of Applied Behavior Analysis*, *43*, 369–380.

Van Houten, R., Malenfant, L., Van Houten, J., & Retting, R. A. (1997). Using auditory pedestrian signals to reduce pedestrian and vehicle conflicts. *Transportation Research Record: Journal of the Transportation Research Board*, *1578*, 20–22.

Van Houten, R., Nau, P. A., Mackenzie-Keating, S. E., Sameoto, D., & Colavecchia, B. (1982). An analysis of some variables influencing the effectiveness of reprimands. *Journal of Applied Behavior Analysis*, *15*, 65–83.

Van Houten, R., Nau, P. A., & Marini, Z. (1980). An analysis of public posting in reducing speeding behavior on an urban highway. *Journal of Applied Behavior Analysis*, *13*, 383–395.

Van Houten, R., Retting, R. A., Farmer, C. M., Van Houten, R. & Malenfant, J. E. L. (2000). Field evaluation of a leading pedestrian interval signal phase at three urban intersections. *Transportation Research Record. No. 1734*, p. 86–91.

Van Houten, R., Rolider, A., Nau, P., Friedman, R., Becker, M., Chalodovsky, I., & Scherer, M. (1985). Large-scale reductions in speeding and accidents in Canada and Israel: A behavioral ecological perspective. *Journal of Applied Behavior Analysis*, *18*, 87–93.

Van Lier, A. C., Van Der Sar, R. M., Muthen, B. O., & Crijen, A. M. (2004). Preventing disruptive behavior in elementary schoolchildren: Impact of a universal classroom-based intervention. *Journal of Counseling and Clinical Psychology*, *72*(3), 467–478.

van Lier, P., Muthen, B., van der Sar, R., & Crijnen, A. (2004) Preventing disruptive behavior in elementary schoolchildren: Impact of a universal classroom-based intervention. *Journal of Consulting and Clinical Psychology*, *72*(3), 467–478.

Van Norman, R. K. (2005). *The effects of functional communication training, choice making, and an adjusting work schedule on problem behavior maintained by negative reinforcement.* Unpublished doctoral dissertation, The Ohio State University, Columbus, OH.

Vannest, K. J., Parker, R. I., Gonen, O., & Adiguzel, T. (2016). *Single case research: Web based calculators for SCR analysis (Version 2.0) [Web-based application]*. College Station, TX: Texas A&M University. Retrieved June 10, 2018. Available from singlecaseresearch.or

Vanselow, N. R., & Bourret, J. C. (2012). Online interactive tutorials for creating graphs with Excel 2007 or 2010. *Behavior Analysis in Practice*, *5*, 40–46. doi:10.1002/jaba.223

Vanselow, N. R., Thompson, R., & Karsina, A. (2011). Data-based decision making: The impact of data variability, training, and context. *Journal of Applied Behavior Analysis*, *44*, 767–780. doi:10.1901/jaba.2011.44-767

VanWagner, M., Van Houten, R., & Betts, B. (2011). The effects of a rectangular rapid-flashing beacon on vehicle speed. *Journal of Applied Behavior Analysis*, *44*, 629–633.

VanWormer, J. J., (2004). Pedometers and brief e-counseling: Increasing physical activity for overweight adults. *Journal of Applied Behavior Analysis*, *37*, 421–425.

Vargas, E. A., Vargas, J. S., & Knapp, T. J. (2017). B. F. Skinner's analysis of verbal behavior: A chronicle (Part I). *Operants*, *3*, 20–23.

Vargas, J. S. (1984). What are your exercises teaching? An analysis of stimulus control in instructional materials. In W. L. Heward, T. E. Heron, D. S. Hill, & J. Trap-Porter (Eds.), *Focus on behavior analysis in education* (pp. 126–141). Columbus, OH: Charles E. Merrill.

Vargas, J. S. (1990). B. F. Skinner–The last few days. *Journal of Applied Behavior Analysis*, *23*, 409–410.

Vasconcelos, M., & Urcuioli, P. J. (2008). Deprivation level and choice in pigeons: A test of within-trial contrast. *Learning & Behavior*, *36*, 12–18.

Vaughan, M. (1989). Rule-governed behavior in behavior analysis: A theoretical and experimental history. In S. C. Hayes (Ed.), *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 97–118). New York: Plenum Press.

Vaughan, M. E., & Michael, J. (1982). Automatic reinforcement: An important but ignored concept. *Behaviorism*, *10*, 217–227.

Vaz, P. C. M., Piazza, C. C., Stewart, V., Volkert, V. M., & Groff, R. A. (2012) Using a chaser to decrease packing in children with feeding disorders. Journal of Applied Behavior Analysis, *45*, 97-105.

Vedora, J., & Barry, T. (2016). The use of picture prompts and prompt delay to teach receptive labeling. *Journal of Applied Behavior Analysis*, *49*, 960–964.

Verriden, A. L., & Roscoe, E. M. (2016). A comparison of preference-assessment methods. *Journal of Applied Behavior Analysis*, *49*, 265–285. doi:10.1002/jaba.302

Vigen, T. (2015). *Spurious correlations*. New York: Hachette Books.

Vladescu, J. C., & Kodak, T. (2010). A review of recent studies on differential reinforcement during skill acquisition in early intervention. *Journal of Applied Behavior Analysis*, *43*, 351–355. http://dx.doi.org/10.1901/jaba.2010.43-351

Vladescu, J. C., & Kodak, T. M. (2013). Increasing instructional efficiency by presenting additional stimuli in learning trials for children with autism spectrum disorders. *Journal of Applied Behavior Analysis*, *46*, 805–816.

Volkert, V. M., & Piazza, C. C. (2012). Pediatric feeding disorders. In P. Sturmey & M. Hersen (Eds.), *Handbook of evidence-based practice in clinical psychology: Vol 1. Child and adolescent disorders* (pp. 323–337). Hoboken, NJ: Wiley. doi:10.1002/9781118156391.ebcp001013

Vollmer, T. R. (1994). The concept of automatic reinforcement: Implications for behavioral research in developmental disabilities. *Research in Developmental Disabilities*, *15*(3), 187–207.

Vollmer, T. R. (2002). Punishment happens: Some comments on Lerman and Vorndran's review. *Journal of Applied Behavior Analysis*, *35*, 469–473.

Vollmer, T. R., & Athens, E. (2011). Developing function-based extinction procedures. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.). (2011). *Handbook of applied behavior analysis* (pp. 317–334). New York: Guilford Press.

Vollmer, T. R., & Hackenberg, T. D. (2001). Reinforcement contingencies and social reinforcement: Some reciprocal relations between basic and applied research. *Journal of Applied Behavior Analysis*, *34*, 241–253.

Vollmer, T. R., & Iwata, B. A. (1991). Establishing operations and reinforcement effects. *Journal of Applied Behavior Analysis*, *24*, 279–291.

Vollmer, T. R., Borrero, J. C., Wright, C. S., Van Camp, C., & Lalli, J. S. (2001). Identifying possible contingencies during descriptive analyses of severe behavior disorders. *Journal of Applied Behavior Analysis*, *34*, 269–287.

Vollmer, T. R., Iwata, B. A., Zarcone, J. R., Smith, R. G., & Mazaleski, J. L. (1993). The role of attention in the treatment of attention-maintained self-injurious behavior: Noncontingent reinforcement and differential reinforcement of other behavior. *Journal of Applied Behavior Analysis*, *26*, 9–21. doi:10.19 01/jaba.1993.26-9

Vollmer, T. R., Joslyn, R., Rubow, C. C., Kronfli, F. R., & Donaldson, J. M. (2016). *The good old Good Behavior Game: Why more studies? Why now?*

Presentation to the 37th annual Berkshire Association for Behavior Analysis and Therapy Conference, University of Massachusetts, Amherst, MA, October 12–14, 2016.

Vollmer, T. R., Marcus, B. A., & LeBlanc, L. A. (1994). Treatment of self-injury and hand mouthing following inconclusive functional analyses. *Journal of Applied Behavior Analysis*, *27*, 331–343.

Vollmer, T. R., Marcus, B. A., Ringdahl, J. E., & Roane, H. S. (1995). Progressing from brief assessments to extended experimental analyses in the evaluation of aberrant behavior. *Journal of Applied Behavior Analysis*, *28*, 561–576. doi:10.1901/jaba.1995.28-561

Vollmer, T. R., Progar, P. R., Lalli, J. S., Van Camp, C. M., Sierp, B. J., Wright . . . Eisenschink, K. J. (1998). Fixed-time schedules attenuate extinction-induced phenomena in the treatment of severe aberrant behavior. *Journal of Applied Behavior Analysis*, *31*, 529–542.

Vollmer, T. R., Roane, H. S., Ringdahl, J. E., & Marcus, B. A. (1999). Evaluating treatment challenges with differential reinforcement of alternative behavior. *Journal of Applied Behavior Analysis*, *32*, 9–23.

Vuchinich, R. E., & Simpson, C. A. (1998). Hyperbolic temporal discounting in social drinkers and problem drinkers. *Experimental and Clinical Psychopharmacology*, *6*, 292–305.

Wacker, D. P., Berg, W. K., Harding, J. W. & Cooper-Brown, L. J. (2011). Functional and structural approaches to behavioral assessment of problem behavior. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 165–181). New York: Guilford Press.

Wacker, D. P., Berg, W. K., Harding, J. W., Derby, M. K., Asmus, J. M., & Healy, A. (1992). Evaluation and long-term treatment of aberrant behavior displayed by young children with disabilities. *Journal of Developmental & Behavioral Pediatrics*, *19*, 235–319.

Wacker, D. P., Harding, J. W., Berg, W. K., Lee, J. F., Schieltz, K. M., Padilla, Y. C., Nevin, J. A., . . . Shahan, T. A. (2011). An evaluation of persistence of treatment effects during long-term treatment of destructive behavior. *Journal of the Experimental Analysis of Behavior*, *96*, 261–282. doi.org/10.1901/jeab.2011.96-261

Wacker, D. P., Lee, J. F., Dalmau, Y. C. P., Kopelman, T. G., Lindgren, S. D., Kuhle, J., . . . Waldron, D. B. (2013). Conducting functional analyses of problem behavior via telehealth. *Journal of Applied Behavior Analysis*, *46*, 31–46.

Wacker, D. P., Schieltz, K. M., Berg, W. K., Harding, J. W., Padilla Dalmau, Y. C., & Lee, J. F. (2017). The long-term effects of functional communication training conducted in young children's home settings. *Education and Treatment of Children*, *40*(1), 43–56.

Wacker, D. P., Steege, M. W., Northup, J., Sasso, G., Berg, W., Reimers, T., et al. (1990). A component analysis of functional communication training across three topographies of severe behavior problems. *Journal of Applied Behavior Analysis*, *23*, 417–429.

Wagaman, J. R., Miltenberger, R. G., & Williams, D. E. (1995). Treatment of a vocal tic by differential reinforcement. *Journal of Behavior Therapy & Experimental Psychiatry*, *26*, 35–39.

Wagaman, J. R., Miltenberger, R. G., & Woods, D. W. (1995). Long-term follow-up of a behavioral treatment for stuttering. *Journal of Applied Behavior Analysis*, *28*, 233–234.

Wahl, E., Hawkins, R. O., Haydon, T., Marsicano, R., & Morrison, J. Q. (2016). Comparing versions of the Good Behavior Game. *Behavior Modification*, *4*, 493–517.

Wahler, R. G., & Fox, J. J. (1981). Setting events in applied behavior analysis: Toward a conceptual and methodological expansion. *Journal of Applied Behavior Analysis*, *14*, 327–338.

Walker, B. D., & Rehfeldt, R. A. (2012). An evaluation of the stimulus equivalence paradigm to teach single-subject design to distance education students via Blackboard. *Journal of Applied Behavior Analysis*, *45*, 329–344.

Walker, B. D., Rehfeldt, R. A., & Ninness, C. (2010). Using the stimulus equivalence paradigm to teach course material in an undergraduate rehabilitation course. *Journal of Applied Behavior Analysis*, *43*, 615–633.

Walker, H. M. (1983). Application of response cost in school settings: Outcomes, issues and recommendations. *Exceptional Education Quarterly*, *3*, 46–55.

Walker, H. M. (1997). *The acting out child: Coping with classroom disruption* (2nd ed.). Longmont, CO: Sopris West.

Walker, H. M., & Buckley, N. K. (1968). The use of positive reinforcement in conditioning attending behavior. *Journal of Applied Behavior Analysis*, *1*, 245–250.

Wallace, I. (1977). Self-control techniques of famous novelists. *Journal of Applied Behavior Analysis*, *10*, 515–525.

Wallace, M. D., & Knights, D. J. (2003). An evaluation of a brief functional analysis format within a vocational setting. *Journal of Applied Behavior Analysis*, *36*, 125–128. doi:10.1901/jaba.2003.36-125

Ward, P., & Carnes, M. (2002). Effects of posting self-set goals on collegiate football players' skill execution during practice and games. *Journal of Applied Behavior Analysis*, *35*, 1–12.

Ward-Horner, J., & Sturmey, P. (2010). Component analyses using single-subject experimental designs: A review. *Journal of Applied Behavior Analysis*, *43*, 685–704. doi:10.1901/jaba.2010.43-685

Warner, S. F. (1992). Facilitating basic vocabulary acquisition with milieu teaching procedures. *Journal of Early Intervention*, *16*, 235–251.

Warren, S. F., Rogers-Warren, A., & Baer, D. M. (1976). The role of offer rates in controlling sharing by young children. *Journal of Applied Behavior Analysis*, *9*, 491–497.

Watkins, N., & Rapp, J. T. (2014). Environmental enrichment and response cost: Immediate and subsequent effects on stereotypy. *Journal of Experimental Analysis of Behavior*, *47*, 186–191.

Watson, D. L., & Tharp, R. G. (2014). *Self-directed behavior: Self-modification for personal adjustment* (10th ed.). Belmont, CA: Wadsworth.

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, *20*, 158–177.

Watson, J. B. (1924). *Behaviorism*. New York: W. W. Norton.

Watson, P. J., & Workman, E. A. (1981). The nonconcurrent multiple baseline across-individuals design: An extension of the traditional multiple baseline design. *Journal of Behavior Therapy and Experimental Psychiatry*, *12*, 257–259.

Watson, T. S. (1996). A prompt plus delayed contingency procedure for reducing bathroom graffiti. *Journal of Applied Behavior Analysis*, *29*, 121–124.

Watson, T. S., Dufrene, B., Weaver, A., Butler, T., & Meeks, C. (2005). Brief antecedent assessment and treatment of tics in the general education classroom: A preliminary investigation. *Behavior Modification*, *29*, 839–857.

Watt, A., Keenan, M., Barnes, D., & Cairns, E. (1991). Social categorization and stimulus equivalence. *The Psychological Record*, *41*, 33–50.

Webber, J., & Scheuermann, B. (1991). Accentuate the positive . . . eliminate the negative. *Teaching Exceptional Children*, *24*(1), 13–19.

Wechsler, D. (2003). *Wechsler intelligence scale for children* (4th ed.). San Antonio, TX: Psychological Corporation.

Weeden, M. A., Mahoney, A., & Poling, A. (2010). Self-injurious behavior and functional analysis: Where are the descriptions of participant protections? *Research in Developmental Disabilities*, *31*, 299–303. doi:10.1016/j.ridd.2009.09.016

Weeks, M., & Gaylord-Ross, R. (1981). Task difficulty and aberrant behavior in severely handicapped students. *Journal of Applied Behavior Analysis*, *14*, 449–463.

Weil, T. M., Hayes, S. C., & Capurro, P. (2011). Establishing a deictic relational repertoire in young children. *The Psychological Record*, *61*, 371–390.

Weiner, H. (1962). Some effects of response cost upon human operant behavior. *Journal of Experimental Analysis of Behavior*, *5*, 201–208.

Weiss, M. J., & Demiri, V. (2011). *Jumpstarting communication skills in children with autism*. Bethesda, MD: Woodbine House, Inc.

Weiss, S. J. (2014). Instrumental and classical conditioning: Intersections, interactions and stimulus. In F. K. McSweeney & E. S. Murphy (Eds.), *The Wiley-Blackwell handbook of operant and classical conditioning* (pp. 417–451). West Sussex, UK: John Wiley & Sons.

Weldy, C. R., Rapp, J. T., & Capocasa, K. (2014). Training staff to implement brief stimulus preference assessments. *Journal of Applied Behavior Analysis*, *47*, 214–218. doi:10.1002/jaba.98

Wenrich, W. W., Dawley, H. H., & General, D. A. (1976). *Self-directed systematic desensitization: A guide for the student, client, and therapist*. Kalamazoo, MI: Behaviordelia.

Werts, M. G., Caldwell, N. K., & Wolery, M. (1996). Peer modeling of response chains: Observational learning by students with disabilities. *Journal of Applied Behavior Analysis*, *29*, 53–66.

West, J., Bergstrom, T., & Bergstrom, C. T. (2010). Big Macs and Eigenfactor scores: Don't let correlation coefficients fool you. *Journal of the American Society for Information Science and Technology*, *61*(9), 1800–1807.

West, R. P., & Smith, T. G. (2002). *Managing the behavior of groups of students in public schools: Clocklights and group contingencies*. Paper presented at The Ohio State University Third Focus on Behavior Analysis in Education Conference, September 21, 2002.

Wetherington, C. L. (1982). Is adjunctive behavior a third class of behavior? *Neuroscience and Biobehavioral Reviews*, *6*, 329–350.

Whaley, D. L., & Malott, R. W. (1971). *Elementary principles of behavior*. Englewood Cliffs, NJ: Prentice Hall.

Whaley, D. L., & Surratt, S. L. (1968). *Attitudes of science*. Kalamazoo, MI: Behaviordelia.

What Works Clearinghouse. (2018). *Standards handbook: Version 4.0*. Washington, DC: Institute of Education Sciences, U.S. Department of Education.

Wheeler, J. J., Baggett, B. A., Fox, J., & Blevins, L. (2006). Treatment integrity: A review of intervention studies conducted with children with autism. *Focus on Autism and Other Developmental Disabilities*, *21*(1), 45–54. doi:10.1177/10883576060210010601

Whelan, R., & Barnes-Holmes, D. (2004). The transformation of consequential functions in accordance with the relational frames of same and opposite. *Journal of the Experimental Analysis of Behavior*, *82*, 177–195.

Whelan, R., Barnes-Holmes, D., & Dymond, S. (2006). The transformation of consequential functions in accordance with the relational frames of more-than and less-than. *Journal of the Experimental Analysis of Behavior*, *86*, 317–335.

White, A., & Bailey, J. (1990). Reducing disruptive behaviors of elementary physical education students with sit and watch. *Journal of Applied Behavior Analysis*, *23*, 353–359.

White, D. M. (1991). *Use of guided notes to promote generalized notetaking behavior of high school students with learning disabilities*. Unpublished master's thesis, The Ohio State University, Columbus, OH.

White, G. D. (1977). The effects of observer presence on the activity level of families. *Journal of Applied Behavior Analysis*, *10*, 734.

White, K. G. (2013). Remembering and forgetting. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. 411–437). Washington, DC: American Psychological Association.

White, M. A. (1975). Natural rates of teacher approval and disapproval in the classroom. *Journal of Applied Behavior Analysis*, *8*, 367–372.

White, T. I. (1993). *Business ethics: A philosophical reader*. New York: Macmillan.

Whitehouse, C. M., Vollmer, T. R., & Colbert, B. (2014). Evaluating the use of computerized stimulus preference assessments in foster care. *Journal of Applied Behavior Analysis*, *47*, 470–484.

Wieseler, N. A., Hanson, R. H., Chamberlain, T. P., & Thompson, T. (1985). Functional taxonomy of stereotypic and self-injurious behavior. *Mental Retardation*, *23*, 230–234.

Wilder, D. A., & Carr, J. E. (1998). Recent advances in the modification of establishing operations to reduce aberrant behavior. *Behavioral Interventions*, *13*, 43–59.

Wilder, D. A., Atwell, J., & Wine, B. (2006). The effects of varying levels of treatment integrity on child compliance during treatment with a three-step prompting procedure. *Journal of Applied Behavior Analysis*, *39*, 369–373. doi:10.1901/jaba.2006.144-05

Wilder, D. A., Majdalany, L., Sturkie, L., & Smeltz, L. (2015). Further evaluation of the high-probability instructional sequence with and without programmed reinforcement. *Journal of Applied Behavior Analysis*, *48*, 511–522. doi:10.1002/jaba.218

Wilder, D. A., Masuda, A., O'Connor, C., & Baham, M. (2001). Brief functional analysis and treatment of bizarre vocalizations in an adult with schizophrenia. *Journal of Applied Behavior Analysis*, *34*, 65–68.

Wilkenfeld, J., Nickel, M., Blakely, E., & Poling, A. (1992). Acquisition of lever press responding in rats with delayed reinforcement: A comparison of three procedures. *Journal of the Experimental Analysis of Behavior Analysis*, *51*, 431–443.

Wilkinson, G. S., & Robertson, G. J. (2017). *Wide Range Achievement Test* (fifth edition). San Antonio, TX: Pearson.

Wilkinson, K. M., & McIlvane, W. J. (2001). Methods for studying symbolic behavior and category formation: Contributions of stimulus equivalence research. *Developmental Review*, *21*, 355–374.

Wilkinson, L. A. (2003). Using behavioral consultation to reduce challenging behavior in the classroom. *Preventing School Failure*, *47*(3), 100–105.

Williams, C. D. (1959). The elimination of tantrum behavior by extinction procedures. *Journal of Abnormal and Social Psychology*, *59*(2), 269–273. doi.org/10.1037/h0046688

Williams, D. E., Kirkpatrick-Sanchez, S., & Iwata, B. A. (1993). A comparison of shock intensity in the treatment of longstanding and severe self-injurious behavior. *Research in Developmental Disabilities*, *14*, 207–219.

Williams, D. L. (2014). Building a theory of Pavlovian conditioning from the inside out. In F. K. McSweeney & E. S. Murphy (Eds.), *The Wiley-Blackwell handbook of operant and classical conditioning* (pp. 27–52). West Sussex, UK: John Wiley & Sons.

Williams, G., Donley, C. R., & Keller, J. W. (2000). Teaching children with autism to ask questions about hidden objects. *Journal of Applied Behavior Analysis*, *33*, 627–630.

Williams, G., Perez-Gonzales, L. A., & Queiroz, A. B. (2005). Using a combined blocking procedure to teach color discrimination to a child with autism. *Journal of Applied Behavior Analysis*, *38*, 555–558.

Williams, J. A., Koegel, R. L., & Egel, A. L. (1981). Response-reinforcer relationships and improved learning in autistic children. *Journal of Applied Behavior Analysis*, *14*, 53–60.

Williams, J. L. (1973). *Operant learning: Procedures for changing behavior*. Monterey, CA: Brooks/Cole.

Wills, H. P., & Mason, B. J. (2014). Implementation of a self-monitoring application to improve on-task behavior: A high-school pilot study. *Journal of Behavioral Education*, *23*, 421–434. doi:10.1007/s10864-014-9204-x

Wilson, A. N., & Gratz, O. H. (2016). Using a progressive ratio schedule of reinforcement as an assessment tool to inform treatment. *Behavior Analysis in Practice*, *9*, 257–260. doi:10.1007/s40617-016-0107-2

Wilson, A. N., Glassford, T. S., & Koerkenmeier, S. M. (2016). Effects of response effort on resurgence. *Behavior Analysis in Practice*, *9*, 174–178. doi:10.1007/s40617-016-0122-3

Wilson, B. A., & Tunnard, C. (2014). Comparison of neuropsychological rehabilitation techniques for unilateral neglect: An ABACADAEAF single-case experimental design. *Neuropsychological Rehabilitation*, *24*(3-4), 382-399.

Wilson, K. G., Hayes, S. C., Gregg, J., & Zettle, R. D. (2001). Psychopathology and psychotherapy. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition*. (pp. 211–237). New York: Plenum.

Wilson, K. M., & Milan, M. A. (1995). Age differences in the formation of equivalence classes. *Journal of Gerontology, 50B*, 212–218.

Windsor, J., Piche, L. M., & Locke, P. A. (1994). Preference testing: A comparison of two presentation methods. *Research in Developmental Disabilities*, *15*, 439–455.

Winett, R. A., & Winkler, R. C. (1972). Current behavior modification in the classroom: Be still, be quiet, be docile. *Journal of Applied Behavior Analysis*, *5*, 499–504.

Winston, A. S., & Baker, J. E. (1985). Behavior analytic studies of creativity: A critical review. *The Behavior Analyst*, *8*, 191–205.

Winterbauer, N. E., & Bouton, M. E. (2010). Mechanisms of resurgence of an extinguished instrumental behavior. *Journal of Experimental Psychology: Animal Behavior Processes*, *36*, 343–353.

Winterbauer, N. E., & Bouton, M. E. (2012). Effects of thinning the rate at which the alternative behavior is reinforced on resurgence of an extinguished instrumental response. *Journal of Experimental Psychology: Animal Behavior Processes*, *38*, 279–291.

Winterbauer, N. E., Lucke, S., & Bouton, M. E. (2013). Some factors modulating the strength of resurgence after extinction of an instrumental behavior. *Learning and Motivation*, *44*, 60–71.

Winterling, V., Dunlap, G., O'Neill, R. (1987). The influence of task variation on the aberrant behavior of autistic students. *Education and Treatment of Children*, *10*, 105–119.

Wirth, O., Slaven, J., & Taylor, M. A. (2014). Interval sampling methods and measurement error: A computer simulation. *Journal of Applied Behavior Analysis*, *47*, 83–100. doi:10.1002/jaba.93

Wiskow, K. M., & Donaldson, J. M. (2016). Evaluation of a lag schedule of reinforcement in a group contingency to promote varied naming of categories items with children. *Journal of Applied Behavior Analysis*, *49*, 472–484.

Witmer, J. F., & Geller, E. S. (1976). Facilitating paper recycling: Effects of prompts, raffles, and contests. *Journal of Applied Behavior Analysis, 9*, 315–322.

Wittgenstein, L. (1953). *Philosophical investigations*. New York: Macmillan.

Wolery, M., & Gast, D. L. (1984). Effective and efficient procedures for the transfer of stimulus control. *Topics in Early Childhood Special Education, 4*, 52–77.

Wolery, M., & Schuster, J. W. (1997). Instructional methods with students who have significant disabilities. *Journal of Special Education, 31*, 61–79.

Wolery, M., Ault, M. J., Gast, D. L., Doyle, P. M., & Griffen, A. K. (1991). Teaching chained tasks in dyads: Acquisition of target and observational behaviors. *The Journal of Special Education, 25*(2), 198–220.

Wolery, M., Lane, K. L., & Common, E. A. (2018). Writing tasks: Literature reviews, research proposals, and final reports. In J. R. Ledford & D. L. Gast (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (3rd ed.) (pp. 43–76). New York: Routledge.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*, 203–214.

Wolf, M. M., Risley, T. R., & Mees, H. L. (1964). Application of operant conditioning procedures to the behavior problems of an autistic child. *Behavior Research and Therapy, 1*, 305–312.

Wolfe, J. B. (1936). Effectiveness of token rewards for chimpanzees. *Comparative Psychological Monographs, 12*, 1–7.

Wolfe, K., & Slocum, T. A. (2015). A comparison of two approaches to training visual analysis of AB graphs. *Journal of Applied Behavior Analysis, 48*, 1–6.

Wolfe, L. H., Heron, T. E., & Goddard, Y. I. (2000). Effects of self-monitoring on the on-task behavior and written language performance of elementary students with learning disabilities. *Journal of Behavioral Education, 10*, 49–73.

Wolfensberger, W. (1972). *The principle of normalization in human services*. Toronto: National Institute on Mental Retardation.

Wolff, R. (1977). Systematic desensitization and negative practice to alter the after-effects of a rape attempt. *Journal of Behavior Therapy and Experimental Psychiatry, 8*, 423–425.

Wolford, T., Alber, S. R., & Heward, W. L. (2001). Teaching middle school students with learning disabilities to recruit peer assistance during cooperative learning group activities. *Learning Disabilities Research & Practice, 16*, 161–173.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press.

Wolpe, J. (1961). The systematic desensitization treatment of neuroses. *The Journal of Nervous and Mental Disease, 132*, 189–203.

Wolpe, J. (1973). *The practice of behavior therapy* (2nd ed.). New York: Pergamon Press.

Wong, S. E., Seroka, P. L., & Ogisi, J. (2000). Effects of a checklist on self-assessment of blood glucose level by a memory-impaired woman with diabetes mellitus. *Journal of Applied Behavior Analysis, 33*, 251–254.

Wood, C. L., Mabry, L. E., Kretlow, A. G., Lo, Y., & Galloway, T. W. (2009). Effects of preprinted response cards on students' participation and off-task behavior in a rural kindergarten classroom. *Rural Special Education Quarterly, 28*(2), 39–47.

Wood, F. H., & Braaten, S. (1983). Developing guidelines for the use of punishing interventions in the schools. *Exceptional Education Quarterly, 3*, 68–75.

Woods, K. E., Luiselli, J. K., & Tomassone, S. (2013). Functional analysis and intervention for chronic rumination. *Journal of Applied Behavior Analysis, 46*, 328–332.

Woolf, S. H. (2008). The meaning of translational research and why it matters. *Journal of the American Medical Association, 299*, 211–213.

Worsdell, A. S., Iwata, B. A., Conners, J., Kahng, S. W., & Thompson, R. H. (2000). Relative influences of establishing operations and reinforcement contingencies on self-injurious behavior during functional analyses. *Journal of Applied Behavior Analysis, 33*, 451–461.

Worsdell, A. S., Iwata, B. A., Dozier, C. L., Johnson, A. D., Neidert, P. L., & Thomason, J. L. (2005). Analysis of response repetition as an error-correction strategy during sight-word reading. *Journal of Applied Behavior Analysis, 38*, 511–527.

Wraikat, R., Sundberg, C. T., & Michael, J. (1991). Topography-based and selection-based verbal behavior: A further comparison. *The Analysis of Verbal Behavior, 9*, 1–17.

Wright, A. N. (2006). The role of modeling and automatic reinforcement in the construction of the passive voice. *The Analysis of Verbal Behavior, 22*, 153–169.

Wright, C. S., & Vollmer, T. R. (2002). Evaluation of a treatment package to reduce rapid eating. *Journal of Applied Behavior Analysis, 35*, 89–93.

Wright-Gallo, G., Higbee, T., Reagon, K., & Davey, B. (2006). Classroom-based functional analysis and intervention for students with emotional/behavioral disorders. *Education and Treatment of Children, 29*(3), 421–436. Retrieved from http://www.jstor.org.libproxy.library.wmich.edu/stable/42899894

Wulfert, E., Greenway, D. E., & Dougher, M. J. (1994). Third-order equivalence classes. *The Psychological Record, 44*, 411–439.

Wunderlich, K. L., & Vollmer, T. R. (2015). Data analysis of response interruption and redirection as a treatment for vocal stereotypy. *Journal of Applied Behavior Analysis, 48*, 749–764.

Wunderlich, K. L., Vollmer, T. R., & Zabala, K. A. (2017). Assessment and treatment of automatically reinforced self-induced emesis. *Journal of Applied Behavior Analysis, 50*, 418–423.

Wyatt v. Stickney, 344 F. Supp. 387, 344 F. Supp. 373 (M.D. Ala 1972), 344 F. Supp. 1341, 325 F. Supp. 781 (M.D. Ala. 1971), aff'd sub nom, Wyatt v. Aderholt, 503 F. 2d. 1305 (5th Cir. 1974).

Wyckoff, L. B. (1959). Toward a quantitative theory of secondary reinforcement. *Psychological Review, 66*, 68–78.

Yanardag, M., Akmanoglu, N., & Yilmaz, I. (2013). The effectiveness of video prompting on teaching aquatic play skills for children with autism. *Disability & Rehabilitation, 35*(1), 47–56. doi:10.3109/09638288.2012.687030

Yanofsky, D. (2015). *It's OK not to start your y-axis at zero*. Retrieved December 5, 2017, from https://qz.com/418083/its-ok-not-to-start-your-y-axis-at-zero/

Yaw, J. S., Skinner, C. H., Skinner, A. L., Maurer, K., Cihak, D., & Wilhoit, B. (2014). Measurement scale influences in the evaluation of sight-word reading interventions. *Journal of Applied Behavior Analysis, 47*, 360–379.

Yell, M. (1994). Timeout and students with behavior disorders: A legal analysis. *Education and Treatment of Children, 17*(3), 293–301.

Yell, M. (1998). *The law and special education*. Upper Saddle River, NJ: Merrill/Prentice Hall.

Yell, M. L., & Drasgow, E. (2000). Litigating a free appropriate public education: The Lovaas hearings and cases. *Journal of Special Education, 33*, 206–215.

Yonkers, K. (2012). *Simple discrimination training in studying stimulus equivalence and math skills acquisition in developmentally-delayed children*. Unpublished master's thesis, University of North Carolina Wilmington.

Yoon, S., & Bennett, G. M. (2000). Effects of stimulus-stimulus pairing procedure on conditioning vocal sounds as reinforcers. *The Analysis of Verbal Behavior, 17*, 75–88.

Young, J. M., Krantz, P. J., McClannahan, L. E., & Poulson, C. L. (1994). Generalized imitation and response-class formation in children with autism. *Journal of Applied Behavior Analysis, 27*, 685–697.

Young, N. D., & Daly, E. J., III (2016). An evaluation of promoting and reinforcement for training visual analysis skills. *Journal of Behavioral Education, 25*, 95–119. doi 10.1007/s10864-015-9234-z

Young, R. K., West, R. P., Smith, D. J., & Morgan, D. P. (1991). *Teaching self-management strategies to adolescents*. Longmont, CO: Sopris West.

Zane, T., Weiss, M. J., Blanco, S., Otte, L., & Southwick, J. (2015). Fads in special education. In R. M. Foxx & J. A. Mulick (Eds.), *Controversial therapies in autism and intellectual disabilities: Fad, fashion, and science in professional practice* (2nd ed.) (pp. 123–135). New York: Routledge.

Zanolli, K., & Daggett, J. (1998). The effects of reinforcement rate on the spontaneous social initiations of socially withdrawn preschoolers. *Journal of Applied Behavior Analysis, 31*, 117–125.

Zentall, T. R., Galizio, M., & Critchfield, T. S. (2002). Categorization, concept learning, and behavior analysis: An introduction. *Journal of the Experimental Analysis of Behavior, 78*, 237–248.

Zerhouni, E. A. (2005). Translational and clinical science—time for a new vision. *Ethics, Medicine, and Public Health*, *2*(2), 256–262.

Zhou, L., Goff, G. A., & Iwata, B. A. (2000). Effects of increased response effort on self-injury and object manipulation as competing responses. *Journal of Applied Behavior Analysis*, *33*, 29–40.

Zhou, L., Iwata, B. A., & Shore, B. A. (2002). Reinforcing efficacy of food on performance during pre- and postmeal sessions. *Journal of Applied Behavior Analysis*, *35*, 411–414.

Zibreg Hargis, D., Pattie, A., Maheady, L., Budin, S., & Rafferty, L. (2017). Using interdependent group contingencies with randomly selected criteria and rewards to enhance homework performance in an eighth-grade classroom. *Journal of Evidence Based Practices for School*, *15*(2), 172–192.

Zimmerman, J., & Ferster, C. B. (1962). Intermittent punishment of S^{Δ} responding in matching-to-sample. *Journal of the Experimental Analysis of Behavior*, *6*, 349–356.

Zinn, T. E., Newland, M. C, & Ritchie, K. E. (2015). The efficiency and efficacy of equivalence-based learning: A randomized controlled trial. *Journal of Applied Behavior Analysis*, *48*, 865–882.

Zlomke, K. R., & Dixon, M. R. (2006). The impact of altering stimulus functions and contextual variables on gambling. *Journal of Applied Behavior Analysis*, *39*, 351–361.

Zuriff, G. E. (1980). Radical behaviorist epistemology. *Psychological Bulletin*, *87*, 337–350.

Zuriff, G. E. (1985). *Behaviorism: A conceptual reconstruction*. New York: Columbia University Press.

Zygmont, D. M., Lazar, R. M., Dube, W. V., & McIlvane, W. J. (1992). Teaching arbitrary matching via sample stimulus-control shaping to young children and mentally retarded individuals: A methodological note. *Journal of the Experimental Analysis of Behavior*, *57*, 109–117.

PHOTO CREDITS

Chapter 19:

Page 504—Silver dollar coin "Somchai Som/Shutterstock"
Page 505—Coin Liberty "Andrey Lobachev/Shutterstock"
Page 505—5 cents coin "Ivan Vdovin/Alamy Stock Photo"
Page 516—American one cent coin "Spiroview Inc/Shutterstock"
Page 536—One dollar note "Steve Stock/Alamy Stock Photo"
Page 536—Five 5 dollar note "Steve Stock/Alamy Stock Photo"
Page 536—20 twenty "Steve Stock/Alamy Stock Photo"
Page 536—Fifty-Dollar-Bill "Onehundred Percent/Alamy Stock Photo"
Page 536—100-dollar bill "imageBROKER/Alamy Stock Photo"

Chapter 30:

Box 30.2a "Elenabsl/123RF"
Box 30.2b "Jim West/Alamy Stock Photo"
Box 30.2c "Peter Bennett/Danita Delimont/Alamy Stock Photo"
Box 30.2d "Alhovik/Shutterstock"
Box 30.2e "Kasezo/123RF"
Box 30.2f "Scott Rothstein/123RF"

NAME INDEX

SUBJECT INDEX

Note: Definitions of terms in italic can also be found in the Glossary.

GLOBAL EDITION

This is a special edition of an established title widely used by colleges and universities throughout the world. Pearson published this exclusive edition for the benefit of students outside the United States and Canada. If you purchased this book within the United States or Canada, you should be aware that it has been imported without the approval of the Publisher or Author.

Applied Behavior Analysis equips students with a sound repertoire of fundamental knowledge about this field that serves as the foundation for more advanced study. The third edition of this pioneering text continues to illustrate behavior analysis principles and procedures with actual applications and real data, helping students acquire the conceptual and technical skills necessary to foster socially adaptive behavior in diverse individuals. The book consists of seven chapters authored by prominent scholars who first reported some of the most significant advances in behavior analysis in publications such as the *Journal of Applied Behavior Analysis, The Analysis of Verbal Behavior*, and the *Journal of the Experimental Analysis of Behavior.*

In addition to updated content and the latest research, this edition includes the following:

- Two new chapters, by noted scholars, on equivalence-based instruction and non-equivalence relations (Chapters 19 and 20) build upon ground-breaking research in the field of applied behavior analysis.
- Multiple-choice and essay-type questions at the end of each chapter test students' mastery of the concepts they have learned.
- Over 2,700 citations and extensive use of quotations and extracts from key publications increase students' exposure to and appreciation for the field's rich primary-source literature.
- More than 150 graphs display original data from peer-reviewed research, often accompanied by detailed descriptions of the procedures used to collect the data represented, helping students develop and refine higher-level visual analysis skills.
- Over 500 technical terms and concepts in the glossary help students finesse the specialized vocabulary of applied behavior analysis.

Pearson

ISBN-10 1-292-32463-5
ISBN-13 978-1-292-32463-0
90000>
9 781292 324630